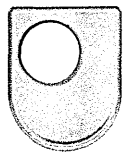
Development of Novel Antitumour Therapeutics Based on DNA Aptamers Against the MUC1/Y Protein

Thesis



DEVELOPMENT OF NOVEL ANTITUMOUR THERAPEUTICS BASED ON DNA APTAMERS AGAINST THE MUC1/Y PROTEIN

BY

VAIDEHI MAKWANA

BSc. MSc.

A THESIS SUBMITTED FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF CHEMISTRY & ANALYTICAL SCIENCES
OPEN UNIVERSITY

OCTOBER 2011

DATE OF SUBMISSION: 7 NOVEMBER 2011
DATE OF AWARD: 25 OCTOBER 2012

ACKNOWLEDGEMENTS

There are many people I would like to take the pleasure in offering my appreciation to for all their support throughout my time spent in researching and developing this project. I would like to thank Antisoma for their sponsorship and my supervisor Dr Sotiris Missailidis, for allowing me to undertake this project and his support during the research.

My sincere appreciation goes to Graham Jeffs, who had provided me with immense assistance during the initial phase of my research and always ensured the perfect maintenance of all the instruments necessary for my research.

I would like to thank Dr Chiara De Pieve for her personal assistance on many aspects of the project and generating good scientific discussions.

My gratitude to Colin Haynes, for providing a great deal of support during my emergencies, last minute requests and always willing to offer his assistance in the setting up of equipment.

I would like to express my appreciation to Julia Barkans and Alejandro Lopez (Department of Biology at the Open University) for accommodating me in their lab. They both also provided me with guidance and their expertise at the required times.

A special thanks goes to Dr Huma Khan for being a great post doc who supported me and gave her personal guidance. She has also been a great friend whose advice and encouragement has been valuable during difficult times.

Last but not least I wish to express my gratitude to my family whom have given me endless support in every way possible during this entire project. Without their encouragement and patience, completion of this thesis would have not been possible.

ABSTRACT

Alterations in the expression of vital proteins of the cellular signalling pathways are at the forefront of molecular abnormalities established in cancer. Therefore, to understand the fundamental mechanisms of cancer initiation and progression, characterisation of tumour markers, such as key proteins like MUC1/Y that are highly expressed in malignant cells, is crucially important.

Rapid drug discovery is a pressing need in the pharmaceutical industry. Agents that are able to bind tightly and selectively to the surface of diseased cells would greatly benefit both disease diagnosis and treatment. An aptamer is a functional oligonucleotide that has the appropriate sequence and structure to form a complex, possessing great affinity and selectivity for its target.

Aptamers of high affinity and specificity have been successfully generated against the MUC1/Y splice variant from a non-conventional single round SELEX system termed SimpLex. The 72 base long aptamers were truncated to an ideal therapeutic length of 25 bases long via two different means. The first entails eliminating the primers and leaving only the central variable region and the second involves truncation of the aptamers based on their predicted secondary structure. Thus, three aptamer versions were taken for further investigation, the complete and two truncated ones.

However, as is it often the case, aptamers lack in vivo stability and are prone to rapid renal filtration and therefore require modification to achieve their full potential *in vivo* as therapeutic agents. Thus, the MUC1/Y selected aptamers were conjugated with a 32 kDa polyethylene glycol (PEG) polymer to increase their molecular weight to *ca.* 40 kDa, an appropriate size to avoid rapid renal clearance.

The truncated aptamers, the pegylated aptamers and the full length aptamers were characterised employing qualitative techniques which include EMSA, affinity column chromatography and DNA thermal denaturation studies. The aptamers were further assessed for their binding, specificity and internalisation by FACs and fluorescence microscopy. The characterisation studies have shown that the anti-MUC1/Y aptamers possess affinity and specificity for the MUC1/Y protein and the conjugation of the aptamers to PEG has further enhanced their binding and internalisation potential. These encouraging results display potential future prospects as therapeutic agents for the pegylated aptamers and warrants their further future evaluation in *in vivo* studies.

LIST OF ABBREVIATIONS

3D ..... Three dimensional
aa ..... Amino acids
ACN ..... Acetonitrile
AMD ..... Age related macular degeneration
APS ..... Ammonium persulfate
bp ..... basepairs
BPH ..... Benign Prostatic Hyperplasia
Cell SELEX ..... Cell-based systematic evolution of ligands by exponential enrichment
CE-SELEX ..... Capillary electrophoresis systematic evolution of ligands by exponential enrichment
cm ..... Centimetre
CT ..... Cytoplasmic tail
Da ..... Daltons
DAG ..... Diacylglycerol
DCM ..... Dichloromethane
DMF ..... Dimethylformamide
DMSO ..... Dimethylsulfoxide
DNA ..... Deoxyribonucleic acid
DS ..... Double Stranded
EDTA ..... Ethylenediaminetetraacetic acid
EGFR ..... Epidermal growth factor receptor
EMSA ..... Electrophoretic mobility shift assay
ER ..... Endoplasmic reticulum
FACs ..... Flow activated cell sorting
FBS ..... Foetal bovine serum
FDA ..... Food and Drug Administration
FITC ..... Fluorescein isothiocyanate
GalNAc ..... *N*-acetylgalactosamine
GlcNAc ..... N-acetylglucosamine
GRB2 ..... Growth factor receptor-bound protein 2
GROs ..... Guanosine-rich oligonucleotides
GSK3β ..... Glycogen synthase kinase 3β
HA ..... Hemagglutinin
HBD ..... Heparin binding domain
HPLC ..... High performance liquid chromatography
hNE ..... Human neutrophil elastase
IFP ..... Interstitial fluid pressure
IgE ..... Immunoglobulin E
LB ..... Luria bertani
Kb ..... Kilobase
Kbp ..... Kilo base pairs
K_D ..... Dissociation constant
KDa ..... Kilo Daltons
K_i ..... Inhibition constant
µA ..... Micro amps
mAbs ..... Monoclonal antibodies
MALDI ..... Matrix-assisted laser desorption/ ionisation
MEM ..... Minimal essential medium eagle
µM ..... Micro Molar
mL ..... Millilitres
mm ..... Millimetres
MP ..... Minimal primer
mM ..... MilliMolar
mRNA ..... messenger RNA
MRT ..... Mean residence time
MWCO ..... Molecular weight cut off
NAALADase ..... *N*-acetyl-α-linked acid dipeptidase
Neu5Ac ..... N-Acetylneuraminic acid

Neu5Gc ..... N-Glycolylneuraminic acid
NHS ..... N-hydroxysuccinimide
NECEEM ..... Non-equilibrium capillary electrophoresis of equilibrium mixtures
NGF ..... Nerve growth factor
NMR ..... Nuclear magnetic resonance
nM ..... Nanomolar
nMol ..... NanoMole
nm ..... Nanometres
nt ..... Nucleotides
NF_KB ..... Nuclear factor $_KB$
PAGE ..... Polyacrylamide gel electrophoresis
pM ..... Picomolar
PM ..... Primer free
PBS ..... Phosphate buffer saline
PCR ..... Polymerase chain reaction
PDGF ..... Platelet derived growth factor
PEG ..... Polyethylene glycol
PSMA ..... Prostate-specific membrane antigen
RES ..... Reticuloendothelial system
RNA ..... Ribonucleic Acid
RNAi ..... RNA interference
RP-HPLC ..... Reverse phase high performance liquid chromatography
RT-PCR ..... Real Time Polymerase chain reaction
SEA ..... Sperm protein, Enterokinase, Agrin
SELEX ..... Systematic evolution of ligands by exponential enrichment
SDS ..... Sodium dodecyl sulphate
SH2 ..... Src Homology
siRNA ..... small interfering RNA
Sos ..... Son of sevenless homology 1
SPR ..... Surface plasmon resonance
SS ..... Single Stranded
T_m ..... Melting temperature
$T_{1/2}$ ..... Half life
TFA ..... Trifluoroacetic acid
TBE ..... Tris Borate-EDTA
TEA ..... Triethylamine
TGF-β2 ..... Transforming growth factor-β2
U ..... Units
UHP ..... Ultra high purity
UV ..... Ultra violet
V ..... Voltage
ValP ..... Valyl phosphate
VEGF ..... Vascular endothelial growth factor
Vis ..... Visible
VNTR ..... Variable number of tandem repeats
WBCT ..... whole blood clotting time

TABLE OF CONTENTS

CHAPTER ONE

INTRODUCTION

1.0 INTRODUCTION

Over the course of many decades, malignancy of different organs has become increasingly prevalent. In many cases, this fatal disease eludes detection, leaving the diagnosed insufficient or no time at all to respond effectively to the therapy administered. It is estimated that 12 million individuals are being diagnosed with this devastating disease, causing mortality in 7.6 million cases every year, equating to approximately one in eight deaths worldwide only from cancer. There are 26 types of cancers of which, cancers of the lung, breast, bowel, stomach, and prostate approximately account for half of all the cancers diagnosed globally. The most common causes of cancer death are cancers of the lung, stomach and liver, whereas breast cancer remains the most prevalent cancer worldwide [1,2].

With cancer on the increase and continuing to be a global public health dilemma, there is a pressing need for the development of novel anticancer therapeutics and effective diagnostic agents. This being the main focus, scientists are evermore researching, following a path of exploration of different methodologies to result in the discovery of innovative and advanced anticancer agents. In order to achieve the design of effective anticancer agents, it is essential to understand the fundamental mechanisms of cancer initiation, progression and characterisation of tumour markers. There are many proteins, growth factors, hormones, antibodies and small molecules which are involved either at the cancer development stage of these processes or during the progression of the malignancy. However, involvement of any of these molecules indicates their potential as tumour biomarkers, such as molecules involved in cell proliferation, initiation of excessive cellular division, and those responsible for cell signalling. Molecules involved in cell signalling are vital as, through mutations, the normal function is bypassed, allowing cell proliferation to occur indefinitely and inhibit apoptosis, which in turn can result in greater expression of particular proteins. Hence, over-expression of proteins in malignant tissues, as opposed to the expression of the proteins found in normal tissues of the same organ, are crucially important and of great interest as potential molecular tumour targeting biomarkers. Following this rationale, a novel target protein, MUC1/Y, a splice variant of the MUC1 glycoprotein, was selected as a potential therapeutic target and as a tumour biomarker for diagnosis in many tumour models.

The therapeutic applications currently available in the market for the treatment of cancer include radiotherapy, chemotherapy and immunotherapy, which employ small chemical molecules, antibodies, peptides, or oligonucleotides. The continuing research into novel anticancer reagents has led to the discovery of a class of nucleic acid-based agents referred to as aptamers. Aptamers are compounds with great potential to bind selectively and with high affinity to their targets, thus possessing the remarkable ability to distinguish between normal and malignant tissues effectively. This, in turn, provides a measureable method for determination of target expression levels within the malignant form. Aptamers are considered to be a rival to antibodies in numerous ways, especially as they pose limited drawbacks with respect to their applications in *in vivo* and *in vitro* studies. Most importantly, the relevance of their use in therapeutics is their ability to reduce or inhibit cell proliferation of cancerous cells or initiate apoptosis via mechanisms of affecting the cellular signalling pathways [3]. Although the MUC1/Y protein has yet not been as widely researched as the native form, MUC1, this isoform has been studied more than many of the other splice variants of MUC1. Hence, the aptamers selected specifically against the MUC1/Y protein can also be used as tools in understanding in greater detail the nature and characteristics of the protein, alongside their use in therapeutic or diagnostic applications.

The major drawback of all aptamers for *in vivo* applications is their degradation by nucleases present in serum and cells, in addition to their rapid elimination from the body. Nevertheless, numerous approaches are available to improve the efficacy of aptamers, which include chemical modifications on the backbone, the sugar or the base moieties of the nucleotide, to prevent nuclease degradation. Modifications which include conjugation of aptamers with high molecular weight polymers, attachment to the phospholipid lamellae and linkage of a multiple number of aptamers of the same sequence to each other, or a combination of these techniques, have a dual functionality of inhibiting nuclease degradation and significantly reducing the swift elimination of aptamers from the body, thus improving their pharmacological properties.

1.1.0 MUCINS

The cells of epithelial organs within the body, such as mammary and salivary glands, respiratory, gastrointestinal, reproductive tracts, kidney and bladder [4], have their luminal surfaces protected by a significantly viscous and slimy secretion, referred to as mucus. Mucus comprises a mixture of components which include water, salts, secreted proteins, mucins and immunoglobulins. Goblet or mucus cells are epithelial cells present on the surface of organs responsible for the secretion of mucus, which serves the purpose of providing a protective layer between the cellular surface and the extracellular environment. Although the mucus mainly operates as a physical selective barrier, the second function is to maintain the cells in their hydrated form via lubrication. From the mixture of components present in the mucus, the main constituents are mucins, which are vastly glycosylated proteins [5].

Mucins are viscous, high molecular weight glycoproteins, appearing in structure to be similar to an elongated rod which can vary significantly in size, resulting in a sequence length from several hundred to a few thousand residues. However, frequently found in most mucins, are regions that are mainly generated from tandem repeat sequences which differ from mucin to mucin with respect to the number, length and amino acid sequence. Nonetheless, despite the differences in size, the majority of the mucin polypeptide chains contain a high ratio of amino acids threonine and serine, whereby the hydroxyl groups of these residues are altered via O-glycosidic bonding with oligosaccharides [6,7]. The amino acid residues can also be linked to oligosaccharides via N-glycosylation, and with the entire carbohydrate composition present on a mucin, this can constitute for up to 90 % of its total molecular weight. From the studies performed on mucins it can be derived that these highly structured glycoproteins have been precisely situated on the cells' surface in order to perform many functions simultaneously. These include guarding the cells from micro-organisms which possess the ability to cause infections via their ability to keep the organisms at a mere distance and capturing them via the carbohydrates' elaborate binding mechanisms. Furthermore, under the umbrella of protective mechanisms, mucins prevent cellular dehydration, shield the cells against physical or chemical injury, whilst allowing particular substances to enter the cell and aiding the passage of various matter along the tracts mentioned earlier [7,8].

There are two categories which mucins are divided into, the secretory and the membrane bound forms. A well founded characteristic of the secreted mucins is their ability to form considerably bulky oligomers resulting from the linking of protein monomers via disulphide bonds. The secreted mucins lack the transmembrane region within the protein and are secreted from the cells into the extracellular space, where they remain at the apical surface of the epithelial cells forming the viscous mucus. Thus, these proteins serve as the first line of defence (Figure 1.0). However, they have not appeared to be associated to any form of tumour initiation or progression [9]. In contrast to the secreted mucins, the membrane bound mucins possess a hydrophobic transmembrane region and a cytoplasmic domain which varies in length. In addition, the membrane bound proteins also maintain their protective function, by extending from the apical surface through their extracellular domain and are incorporated in the formation of the mucus gel. Thus, they operate as the second line of defence and operate as sensors, detecting any external disturbances in the environment. Any alterations in the external environment, such as pH and ionic composition or interaction of any foreign substances binding to the extracellular domain can lead to conformational changes, promoting a signalling cascade within the cell [10].

From the family of mucins, the glycoproteins identified under the two categories include eight secreted mucins (MUC2, MUC5AC, MUC5B, MUC6, MUC7, MUC8, MUC9 and MUC19) [11] and ten membrane bound mucins (MUC1, MUC3A, MUC3B, MUC4, MUC12, MUC13, MUC15, MUC16, MUC17, and MUC20) [10].

1.2.0 MUC1

Of the many membrane bound proteins, the protein commonly referred to as MUC1 is the first mucin from the many to be cloned and have its full cDNA and genomic clones identified for both human and mouse. Hence, the in depth study of this protein has unveiled its numerous properties and functions, which are crucial for many precise biological roles. The entire cDNA sequence of the human MUC1 gene encodes the full protein consisting of three distinctive domains, the extracellular domain, transmembrane and cytoplasmic domains [6].

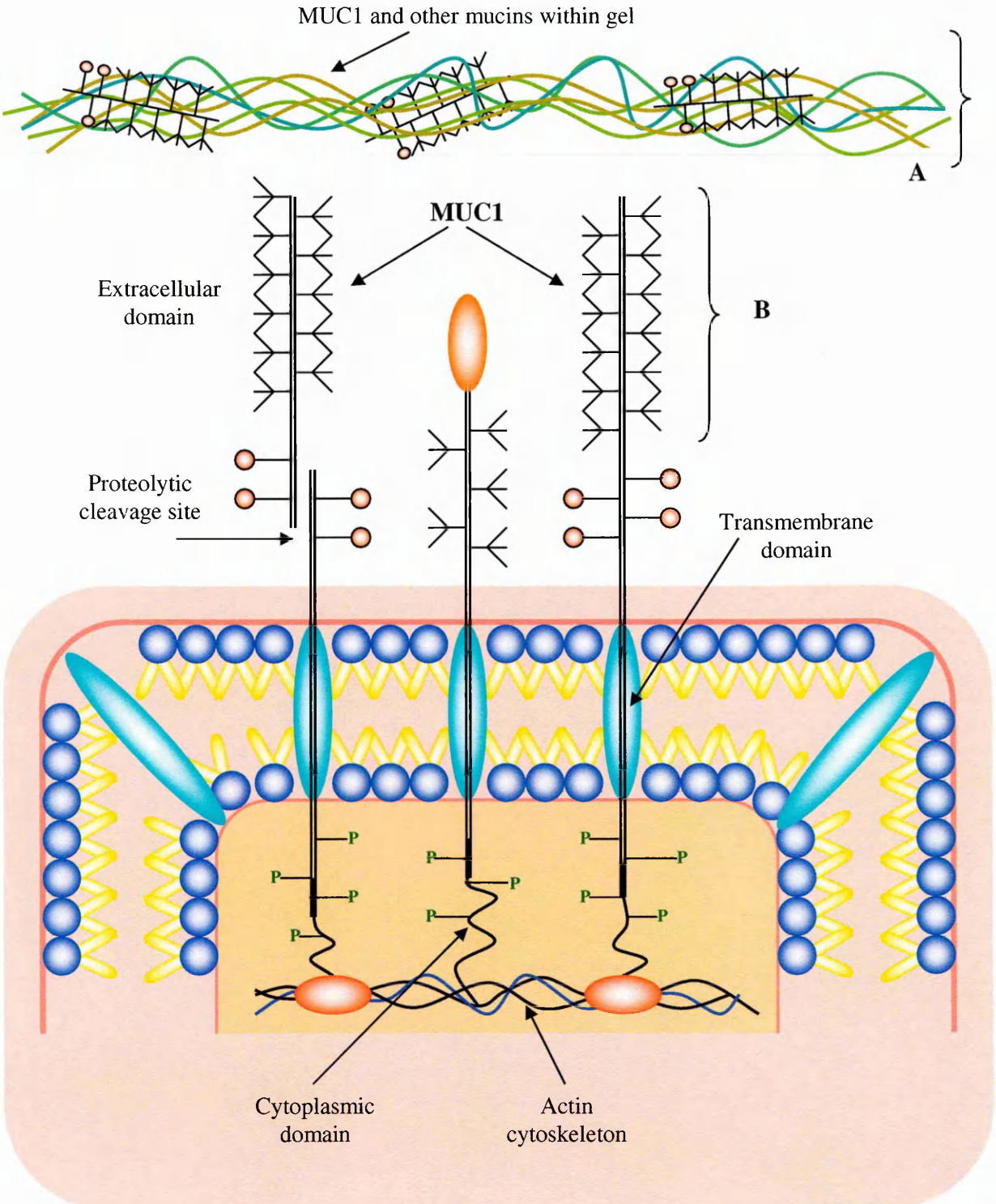


Figure 1.0. Illustration of a mucosal cell surface and membrane protected by mucins. (**A**) First line of defence: the mucus gel, consisting of secreted or shed mucins. (**B**) Second line of defence: MUC1 is present on the apical surface of cells as a heterodimer that can be released by shedding or proteolysis at the proteolytic cleavage site shown. The release can potentially initiate signalling within the cell due to changes in the extracellular environment. Adapted from ref [9,12]

 = O-linked carbohydrate,  = N-linked carbohydrate.

1.2.1 MUC1 GENE & PROTEIN STRUCTURE

The MUC1 gene produces many MUC1 protein variants, of which the major product that has been studied in considerable depth is the MUC1 mucin, a polymorphic transmembrane glycoprotein. The MUC1 gene consists of 7 exons with a region of an estimated length of 4 to 7kb of DNA, although this can vary with respect to the number of 60bp tandem repeat units present within the variable number of tandem repeats (VNTR) domain situated in the exon 2 [13,14].

The MUC1 protein is comprised of three domains: a substantial proportion of the protein is encompassed by the extracellular domain, followed by a transmembrane domain and a short cytoplasmic domain (Fig. 1.0). The extracellular domain is a major part of the core protein as a result of 20 – 125 tandem repeats of 20 amino acid sequences, whereby each 20 amino acid repeat is rich in serine, threonine, proline, alanine and glycine residues. The serine and threonine resides within the tandem repeat sequence, provide five potential *O*-glycosylation sites, hence attributing up to 50% of the molecular mass to the carbohydrates present on the protein. Amino acids alanines and glycines are considered to aid the action of glycosylation enzymes in close proximity. The proline residues are precisely situated between the glycosylation sites, separating the regions of the attached carbohydrates providing an inflexible construction to the extended protein. This large domain protrudes 200 to 500 nm above the membrane, extending out far beyond the cell surface in comparison to the other membrane-bound proteins. Due to the large distance spanning of the domain, along with the high number of proline residues, this has encouraged the likely formation of polyproline β-turn helixes, which are stabilised by the attached oligosaccharides [15,16]. A hydrophobic transmembrane domain, connected to the extracellular domain, is a stretch of 31 amino acids, which links to the phosphorylated cytoplasmic tail consisting of 72 amino acids [17]. The cytoplasmic domain has appeared to be networked with the actin cytoskeleton and is proposed to be involved in signalling pathways [6,18]. The three domains of the core protein has an approximate mass of 120 – 225 kDa, however, the glycosylated mature form of the protein can have an increased weight in the range of 240 – 450 kDa [13].

MUC1 has also been referred to as MUC1/REP, MUC1/TM, Episialin, H23Ag, CA15-3, ETA (Epithelial Tumour Antigen), PEM (Polymorphic Epithelial Mucin), EMA (Epithelial Membrane Antigen), MCA (Mammary Carcinoma Antigen) [17], Epitectin, PAS-0, DF3 and NCRC11. Many of these MUC1 designations have evolved from antibody and antigen classifications, though PAS refers to a positive indication when utilising a stain, periodic acid-Schiff's reagent for carbohydrates, whereas episialin signifies the high presence of sialic acid on epithelial glycoproteins [12].

1.2.2 PROTEOLYTIC CLEAVAGE & SOLUBLE MUC1

Although MUC1 is typically a transmembrane protein, a soluble form of MUC1 has been detected in tissue culture supernatants and bodily fluids. This has suggested shedding of the extracellular domain from the MUC1/REP or a direct secretion of a MUC1 variant, MUC1/SEC, within the soluble form. The soluble form of MUC1 lacks both the transmembrane domain and the cytoplasmic tail proposing that the soluble form is released from the membrane by proteases via two potential mechanisms. Upon translation of the MUC1 protein, a proteolytic cleavage of the extracellular domain occurs within the endoplasmic reticulum resulting in two subunits, which then progresses to the cell surface as a heterodimer complex [19,20]. The two subunits generated, N-terminal and C-terminal, also referred to as α and β [21] respectively, interact with each other specifically via strong non-covalent binding, although the bonds are susceptible to sodium dodecyl sulphate (SDS). The stability of this interaction has been demonstrated by the heterodimer remaining intact despite the several recycling steps of the protein re-entering the cell to the Golgi apparatus for further glycosylation and proceeding back to the cell surface [19,22]. The potential proteolytic cleavage site appears to be within a sequence of 18 amino acids, in between 71 to 53 amino acids upstream of the transmembrane domain, where the N-terminal subunit is the product consisting of the larger part of the extracellular domain, including the variable number of tandem repeats (VNTR). The smaller segment (20-30kDa) of the extracellular domain consists of the remaining 58 amino acids, along with the transmembrane and cytoplasmic domain is positioned at the C-terminal [14,19]. It has been suggested from various studies that the larger portion of the extracellular domain is released from the heterodimeric complex via possible mechanisms not yet clearly identified, although two possibilities

have been proposed; dissociation of the N-terminal subunit from the C-terminal subunit or a second proteolytic cleavage. This release of the N-terminal may be a possible means to signal to the cell interior of an external change in the environment [19,23]. Research has also led to the understanding that the proteolytic cleavage site remains unaffected by any number of tandem repeats which are subjected to heavy glycosylation. Furthermore, it has been observed from experimental results that proteolytic cleavage and arrangement of the heterodimeric complex not only occurs in normal epithelial cells but also within malignant cells [20]. In addition, similar to the MUC1 soluble forms, MUC1/SEC also lacks the transmembrane and the cytoplasmic domain. However, MUC1/SEC can be differentiated as this splice variant possess the potential to be secreted directly from the cell [14].

1.2.3 PROPERTIES & FUNCTION OF MUC1

MUC1 possess many properties that serve its numerous functionalities, which include preventing dehydration and protection of the cell surface from foreign material. Additionally, it possesses a dual functionality of performing as an adhesive and an anti-adhesive molecule. The oligosaccharides present on the protein prevent moisture from escaping by binding to water molecules, thus keeping the moisture in close vicinity of the cell surface and permitting the cells to remain hydrated. Furthermore, they play a significant role in the formation of the mucus gel and provide lubrication. The extended extracellular domain of MUC1, with a network of oligosaccharides, aids in protecting the cell surface from direct contact with the external environment via utilisation of two main pathways. Oligosaccharides either provide a steric obstruction, by blocking the entry of any unrecognised material, or capture bacterial organisms through some form of interaction. Upon any interaction of foreign materials with the protein, cellular changes are promoted by means of activating greater tyrosine phosphorylation on the cytoplasmic tail, indicating to the interior of the cell, the presence of abnormal activities in its external environment. In order to terminate the extension of each carbohydrate chain, the chain is either capped at the end with sialic acid, fucose or sulphate, attributing a negative charge on the molecule. There are several functionalities associated with the presence of sialic acids which determine, depending on the expression of the sialic acid, the fate of cell-cell interactions and cell-matrix formation [24].

A change in cell-cell and cell-matrix interactions is observed through a greater expression of sialic acids causing steric hindrance brought about from the negative charge. As the number of tandem repeats is reduced, so is the *O*-glycosylation, leading to increased levels of sialyl transferases prompting greater sialylation. The presence of sialic acid (Figure 1.1) in normal, non-cancerous cells, demonstrates two functions. Firstly, sialic acid allows the lumens of the vessels to remain open through strong repulsion generated from the negative charge. Secondly, they have the capability to bind to foreign microorganisms preventing them from reaching the cell surface and causing any damage [12]. The degree of sialylation appears to increase each time with the recycling process of the MUC1 protein. However, this can result in the down regulation of the immune response, as access of antibodies binding to their corresponding peptide regions is limited through either steric hindrance or masking of the antigen present on the tumour-associated MUC1. With greater expression of sialic acid, MUC1 performs as an anti-adhesive molecule, preventing cell-cell interactions and cell-matrix formation due to the negative charge repulsion, as well as providing rigidity to the protein. As MUC1 is typically expressed only on the apical surface, it is unable to interfere with the adhesion molecules situated on the basolateral side. MUC1 however, is able to inhibit interaction with proteins present on the opposite side of the cells, preventing non-essential adhesive interactions. Furthermore, in the presence of lectin molecules, which are sugar binding proteins, cell-cell interactions can take place by interacting with the oligosaccharides on the extended extracellular domain [16,25].

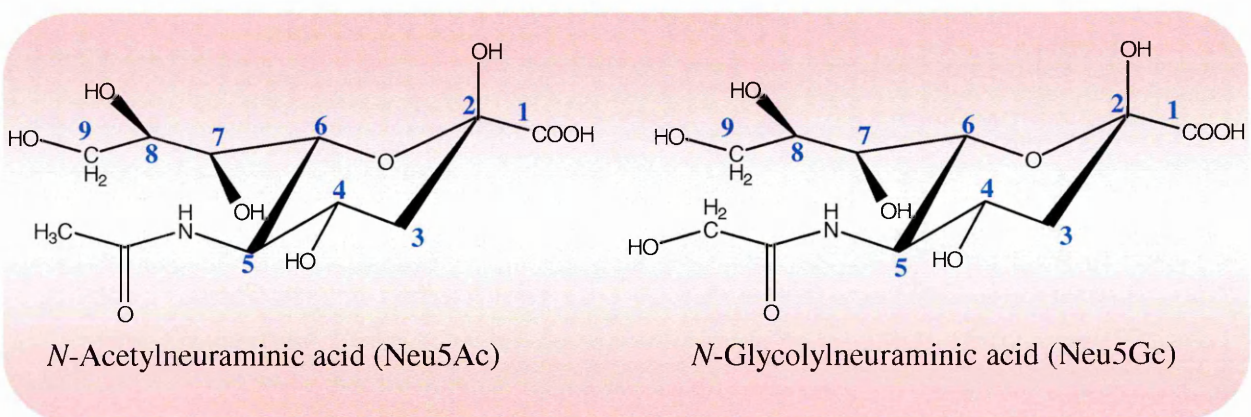


Figure 1.1. Structure of the two major sialic acid derivatives. Neu5Ac is the most abundant sialic acid in all living systems. Neu5Gc is commonly disseminated in the mammalian system, excluding healthy humans. However, the presence of Neu5Gc can be found on malignant cells. Adapted from ref [26].

As the number of tandem repeats decrease, resulting in a considerable reduction of *O*-glycosylation, carbohydrates such as sialyl Lewis^x and sialyl Lewis^a (Figure 1.2) are significantly increased on MUC1 expressed by tumour cells. Sialyl Lewis^x and sialyl Lewis^a extensively promote the anti-adhesive properties of the protein and have an important role in advancing metastasis. Alternatively, sialyl Lewis carbohydrates can also interact with adhesion molecules like P- and E-selectin. These selectin molecules which bind to leukocytes and endothelial cells, forms a complex with the sialyl Lewis carbohydrates present on tumour cells and leukocytes, whereby this interaction can lead to extravasation. Further aiding the metastatic process is the presence of the soluble MUC1 expressed with excessive sialyl Lewis carbohydrates, thus upon binding to the selectin molecules, the leukocytes are distracted away from the tumour site [24,25].

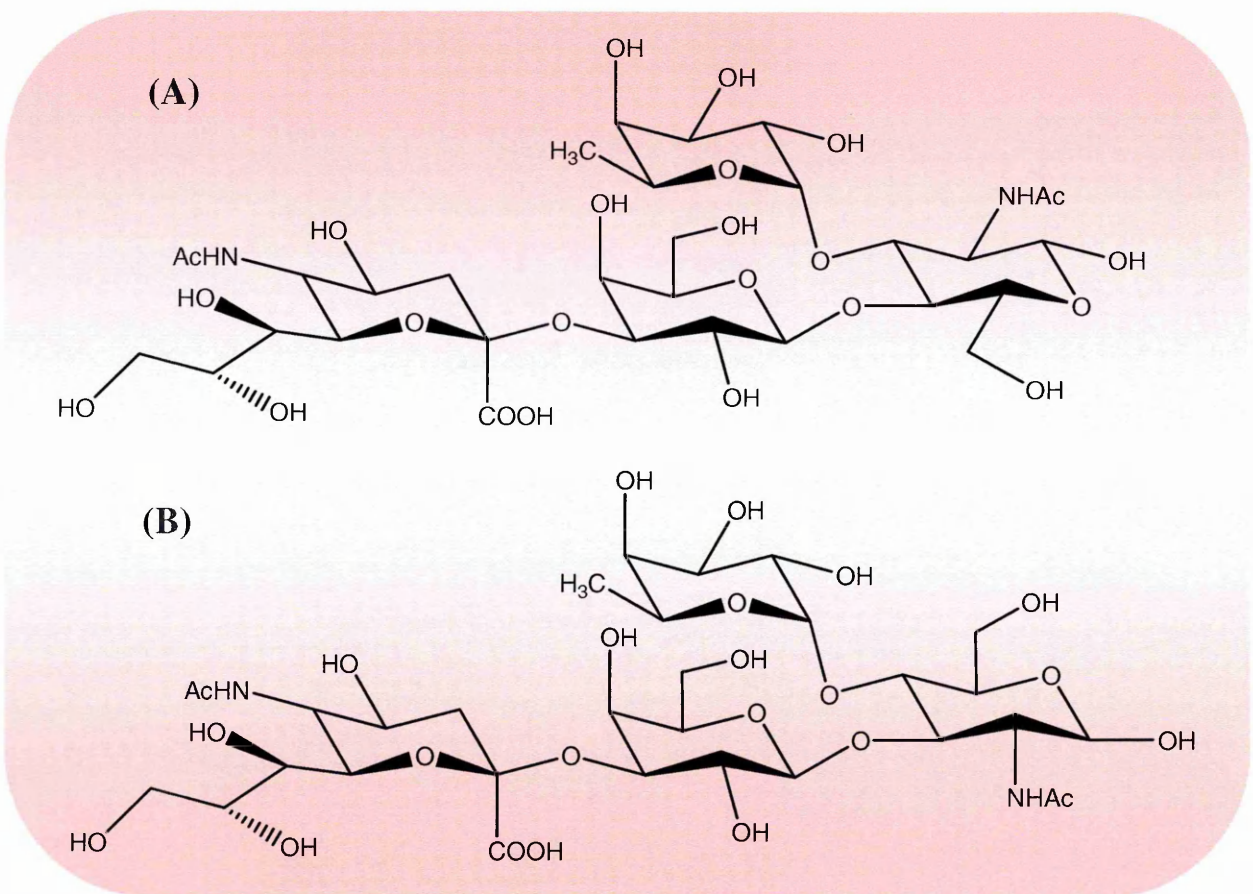


Figure 1.2. Structure of Sialyl Lewis carbohydrates. **(A)** Structure of sialyl Lewis^a and **(B)** sialyl Lewis^x. Adapted from ref [27].

Other than the adhesive and non-adhesive, glycosylated and under glycosylated properties, MUC1 is also associated with immune regulation and/or immune suppression, which is further explained in the following section, 1.2.4.

1.2.4 EXPRESSION OF MUC1

Quantitatively, the expression of each type of mucin is typically varied to some degree across the numerous different organs and tissues. However, aside from the fluctuating levels of mucin expression, the MUC1 protein also possess the ability to be expressed in different morphological structures as a result of differing VNTR regions and glycosylation patterns. The variations may possibly be dependent upon the functional role required to perform a biological process [12]. Expression of the MUC1 mucin, under normal conditions, is restricted to the apical surface on most epithelial membranes. However, beyond the last decade, expression of MUC1 on hematopoietic stem cells, which are stem cells that give rise to all varying type of blood cells, has also been discovered [9,28]. This appears to suggest that MUC1 may have an immunomodulatory role in the functioning of the immune system [29,30]. In particular, in organs dependent upon steroidal hormones for the regulation of cell proliferation, for example in the mammary gland where alternating cycles of active and inactive periods take place, the expression of MUC1 is seemingly elevated during the cell multiplication phase [12]. Within the endometrium, the mucous membrane lining the uterus, a similar trend is also observed, whereby during a particular phase of the menstrual cycle, which is progesterone dominated, greater expression of MUC1 is observed. This progesterone regulated expression of the protein has resulted in either a reduced or inhibited T-cell stimulation within the reproductive tract, which has been found to permit the implantation of the foetus. Aside from the fluctuating levels in the expression of MUC1 within the endometrium, the degree of glycosylation on the protein appears to be much varied. As it appears, MUC1 on normal epithelial surfaces is more densely glycosylated in comparison to the protein found in the endometrium and the serum of expectant mothers, where the sugar chains are significantly truncated [14]. The abnormal form of glycosylation has been found to initiate an immune system response, by activating the production of anti-MUC1 T cells. Thus, it has been proposed that this in-built mechanism performing as a natural immuniser may act as a barrier against the formation of breast cancer during the pregnancy period [29,31].

From the results and observations obtained through numerous investigations performed on the expression of MUC1, it can be postulated that under the normal biological organisation, MUC1 can operate as a regulator of the immune system or an immune suppressor depending upon the role that is required by the protein. However, the precise mechanism of how MUC1 operates in the immune regulatory system is yet to be determined [29].

1.2.5 ROLE OF MUC1 IN SIGNAL TRANSDUCTION

The MUC1 transmembrane protein has not been designated to a particular function, though the protein appears to be responsible for numerous activities, as previously mentioned, including detecting changes within the extracellular environment, protection of cells, cellular adhesion and signal transduction. It has been demonstrated that multiple functional properties of MUC1 can come into action simultaneously via the protein's elaborate structure. The extracellular domain of MUC1 responds to any external change in the environment and this is communicated to the interior of the cell by the means of heterodimer dissociation or by alteration of the domain through some form of binding. The intracellular region of MUC1 provides a platform to interact with different kinds of signalling proteins [11] (GRB2 / Sos / Ras [32], ErbB receptors [33], β-Catenin [34], c-Src tyrosine kinase [35]), whereby upon initiation of the external stimuli, signals appear to be carried to the nucleus via the interaction between the cytoplasmic tail (CT) of MUC1 and signal transducing agents. The cytoplasmic region of MUC1 possesses seven tyrosine residues, which provide several moieties that represent varying degree of binding potential to different forms of signal transducing agents. The MUC1 CT is composed of 72 amino acids of which seven are tyrosine residues, where phosphorylation of these residues has first been demonstrated by Zrihan-Licht et al [17] who has strongly suggested that the MUC1 CT is involved in signal transduction. Amongst the seven tyrosine residues, four tyrosines are positioned in an eloquent manner within the entire amino acid sequence, whereby upon phosphorylation they are projected as signalling receptors and have been noted to allow MUC1 to function in a similar fashion to that of cytokine receptors. Evidence from extensive studies has established that these four tyrosine residues are phosphorylated in response to extracellular stimuli and each of them is able to interact with the Src homology (SH2) domain belonging to different

protein kinases. The SH2 domain is a protein domain which is structurally conserved, found within many signal transducing proteins, which assist proteins in locating other proteins by recognising phosphotyrosine residues. This suggests that each phosphotyrosine residue contributes to a particular functional role and each maybe be accountable for responding differently to the extracellular stimuli [36].

The phosphorylated tyrosine residues of the MUC1 CT have been shown to interact with adaptor proteins such as growth factor receptor-bound protein 2 (GRB2). GRB2 is associated with growth factor receptors involved in signalling pathways [37]. Pandey et al [32] has also shown the binding between the cytoplasmic tail and GRB2 in breast cancer cells. However, results have also shown that upon formation of this complex a further interaction takes place with the son of sevenless homology 1 (Sos) protein. The GRB2, whilst in complex with the MUC1 CT, is also able to bind via the SH3 domain to the Sos, forming a MUC1-GRB2-Sos complex. This, in turn, interacts with the Ras protein and is responsible for communicating signals from the exterior of the cell to the nucleus.

Another interaction between the intracellular region of MUC1 and β-catenin has also been observed [34]. However, in this case a protein kinase referred to as c-Src has been found to phosphorylate the tyrosine residues present in a particular place within the entire sequence. Subsequent to phosphorylation, the SH2 domain of Src is able to directly interact with that particular phosphotyrosine residue to prevent the binding of MUC1 with glycogen synthase kinase 3β (GSK3β) and instead promote the interaction with β-catenin. Upon the formation of MUC1- β-catenin complex, the complex is transported to the nucleus, where β-catenin is responsible for the increase in the expression of cell cycle progression genes [38]. Through this mechanism, MUC1 is able to increase the levels of β-catenin present in the cytoplasm and nucleus, resulting in a continuation of the inhibition of phosphorylation by GSK3β. However, the formation of the MUC1- GSK3β complex can inhibit the MUC1- β-catenin formation, affecting cell adhesion and resulting in decreased MUC1-associated tumorigenicity [10,35,38].

A further participation of MUC1 in signal transduction has been demonstrated by MUC1 expressed on malignant cells, through its ability to interact with carbohydrate-binding proteins via the carboxy

terminal. The complex formed upon the interaction between the carbohydrate binding protein known as epidermal growth factor receptor (EGFR) and MUC1 mediates signalling pathways that have been shown to contribute in cancer cell growth, promoting tumour growth formation and metastasis [33,39,40]. The prevention of the MUC1 and EGFR interaction has resulted in a decrease in proliferation, migration and invasion [40,41]. The inhibition of this type of interaction can take place via the natural regulatory system, whereby other proteins are able to phosphorylate and bind to the phosphotyrosine residues in the CT. On the other hand, if the protein responsible for phosphorylating the residues does not bind to the CT, then the protein can mediate the binding with further relevant proteins [40].

In addition to the importance of the cytoplasmic tail present in MUC1 for signal transduction, Kohlgraf et al [42] have also demonstrated the impact of the presence of the tandem repeats. The lack of either the tandem repeats or the cytoplasmic tail has resulted in the over expression of MUC1 and the cells have had an amplified tendency to invade lymph vessels and metastasise. If the primary functions of MUC1 such as sensory, intracellular signalling and adhesion properties are dependent upon the tandem repeats and the CT, then the absence of either of the two domains, which leads to aggressive characteristics of the overexpressed MUC1, could support the hypothesis that the normal properties and function of the protein are not maintained. This has been demonstrated by the altered form of MUC1 expressed in cancer cells. Following on from this, it has been demonstrated that excessive expression of the full length MUC1 protein has influenced the expression levels of other proteins, thus effecting the invasion and degradation of the extracellular matrix. However, this cannot be achieved in the absence of either the tandem repeats or the cytoplasmic tail. Furthermore, the signals initiated are dependent upon the different types of interaction MUC1 would have. This mainly relies on where the expression of MUC1 has taken place on the cell surface and if the appropriate ligands are present within close proximity of MUC1 for an interaction to take place.

In brief, only a few selected possible signalling pathways involving the extracellular and intracellular regions of the MUC1 protein and interacting proteins have been described above. From this it can be concluded that the intracellular processes and signalling pathways involve an extremely complex

organisation. MUC1 is considered to possess versatile attributes through its involvement in many different physiological roles such as cell growth and differentiation, cell-cell and cell-matrix adhesion, oncogenesis, and immunity. However, MUC1 also functions in signal transduction through the presence of a variety of binding motifs on the cytoplasmic tail. This increases the availability and possibility of interaction with signalling proteins, permitting the formation of a comprehensive network of molecular interactions.

1.2.6 GLYCOSYLATION

The core of the MUC1 protein, once processed in the endoplasmic reticulum (ER) undergoes proteolytic cleavage within the ectodomain region, which initially comes from the transmembrane domain and these two portions remain intact, forming a heterodimeric complex. At the extracellular portion of MUC1, the tandem repeats are the main feature of the protein core. This domain accommodates particular amino acids which are appropriately positioned for the potential attachment of glycans. The oligosaccharide constructs present on the protein core are found to be in an array of diverse forms, which is determined by the type of tissue the protein is expressed upon and how the glycosyltransferases are expressed within that particular tissue [9,43]. The glycosylation process is a post-translational development which involves two types of glycosylation, the *N*- linked and the *O*- linked glycosylation. The section within the MUC1 extracellular domain between the beginning of the transmembrane domain and the VNTR region includes five potential sites for *N*-glycosylation and within the VNTR each tandem repeat is bearing five potential *O*-glycosylation sites [44].

A considerable amount of carbohydrate content on the extracellular portion of the protein is owing to the attachment of the *O*-glycans through covalent links with the hydroxyl groups of serine and threonine residues within the tandem repeats. In *O*-glycosylation, the process of adding the sugars, *N*-acetylgalactosamines (GalNAc), occurs by attaching each one individually and in sequential order as the protein passes through the Golgi, with each addition being catalysed by glycosyltransferases. Once the GalNAc residues have been attached along the core of the extracellular domain, the basis for the addition of numerous other sugars has been created. Hence, from each of these core sugars an extension can be

generated by further attaching N-acetylglucosamine (GlcNAc) and galactose in an alternating fashion, resulting in the formation of side chains which can be either straight or branched. The actual composition of the oligosaccharide chains can vary with individual protein cores even if the protein is expressed on the same tissue. This closely depends on the level of activity of the specific glycosyltransferases available and the position of the protein as it passes through the Golgi with relation to other enzymes that are present in the vicinity that can also compete for the same substrate. The cessation of the chain extension procedure is achieved by the addition of terminal sugars, such as GalNAc, GlcNAc, galactose, fucose, sialic acid or by sulphation. The availability of the number of enzymes responsible for adding sialic acid can significantly influence the length of each *O*-glycan chain, as addition of the sialic acid during the early stages of *O*-glycosylation can inhibit the chain extension process and affect the final carbohydrate structure [45,46]. Although the final termination of the chain extension by adding sialic acid residues is performed in the Golgi, the process is not complete through the first round, as at the end of the first round, the form that is exposed on the plasma membrane is not the mature form. Subsequent to the exposure of the protein to the plasma membrane, MUC1 is reinternalised via endocytosis for greater sialylation in the Golgi and finally the recycled mature form of the protein is brought to the cell surface. A number of recycling rounds are required in order to fully sialylate MUC1, whilst the shorter lengths of the *O*-glycan present allow internalisation at a faster pace than the mucin bearing considerably longer lengths of the *O*-glycans [25,44,47].

The other form of glycosylation which is significantly different and has not been studied as extensively as the *O*-glycosylation is the *N*-glycosylation. Unlike the *O*-glycans, which are moderately small residues that can be attached to adjacent amino acids on the protein core, *N*-glycosylation cannot occur in this manner. The *N*-glycans are large oligosaccharides which cannot be packed closely, hence are required to be situated in a spacious manner [48]. The addition of *N*-glycans also appears to contribute towards the functional properties of MUC1 and is particularly important for the stability, folding, transport, membrane trafficking and secretion for these type of glycoproteins. *N*-glycosylation also contributes significantly to the particular role of MUC1 in the apical expression and sorting of glycoproteins present amongst polarised cells. For example, if the addition of *O*-glycan remains incomplete and an early onset

of sialylation has occurred, this would result in modification and perhaps greater polarisation in the protein, as opposed to the normal form. Thus, it would generate two forms of the protein, consequently leading to the sorting of the two forms. For *N*-glycosylation to come about, the *N*-glycans are attached to the asparagine residues present in the protein core. This type of attachment takes place in the ER subsequent to protein synthesis. The oligosaccharides, which are bound to the asparagine residue, are a mannose rich *N*-glycans, $Glc_3Man_9GlcAc_2$, which are initially trimmed by mannosidases in the Golgi, followed by further enzymatic addition of other saccharides. Furthermore, the procedure of recycling the protein through the Golgi permits the continuation of processing the high mannose glycans, which is observed through the different profiles of the *N*-glycosylation [47,49].

Both forms of glycosylation on the MUC1 protein contribute towards the overall properties and function of the protein. The highly glycosylated extracellular domain plays a role in bulking up the weight of the glycoprotein, with the addition of providing stiffness to its overall structure. The *O*-glycosylation also takes part in cell protection and cell adhesion, whilst the *N*-glycosylation is involved in further intricate biological roles such as transport, secretion and membrane trafficking.

1.3.0 CANCER ASSOCIATED MUC1

The epithelial mucin MUC1 is a functional protein with important roles in the survival of the cell. However, MUC1 is also distinctly associated with many malignancies, where several of the original characteristics of the protein have been lost. As opposed to the typical expression, which is restricted to the apical surface of normal epithelial cells, MUC1 has notably been observed to be overexpressed in many adenocarcinomas such as those of the breast, lung, ovary, pancreas, prostate and numerous other epithelial organs. Other cancers also expressing MUC1 to a substantial degree include multiple myeloma, lymphoma and particular leukaemias [8]. Greater expression is due to the loss of polarisation. However, MUC1 on tumour cells displays many different features compared to the native protein, such as significantly reduced glycosylation due to the lessening in the number of tandem repeats present in the extracellular domain, combined with the extra addition of sialic acids which also distorts the cellular adhesion properties.

On malignant cells, it has been clearly demonstrated that the carbohydrates are unevenly distributed along the domain and branches of the oligosaccharides are drastically truncated (Figure 1.3), with chain lengths composed of only one to six sugar units. As a consequence of this, although the polarised composition from the oligosaccharides has been lost, allowing the expression of MUC1 to take place all over the cells surface, larger portions of the malignant MUC1 is exposed in the extracellular environment [29].

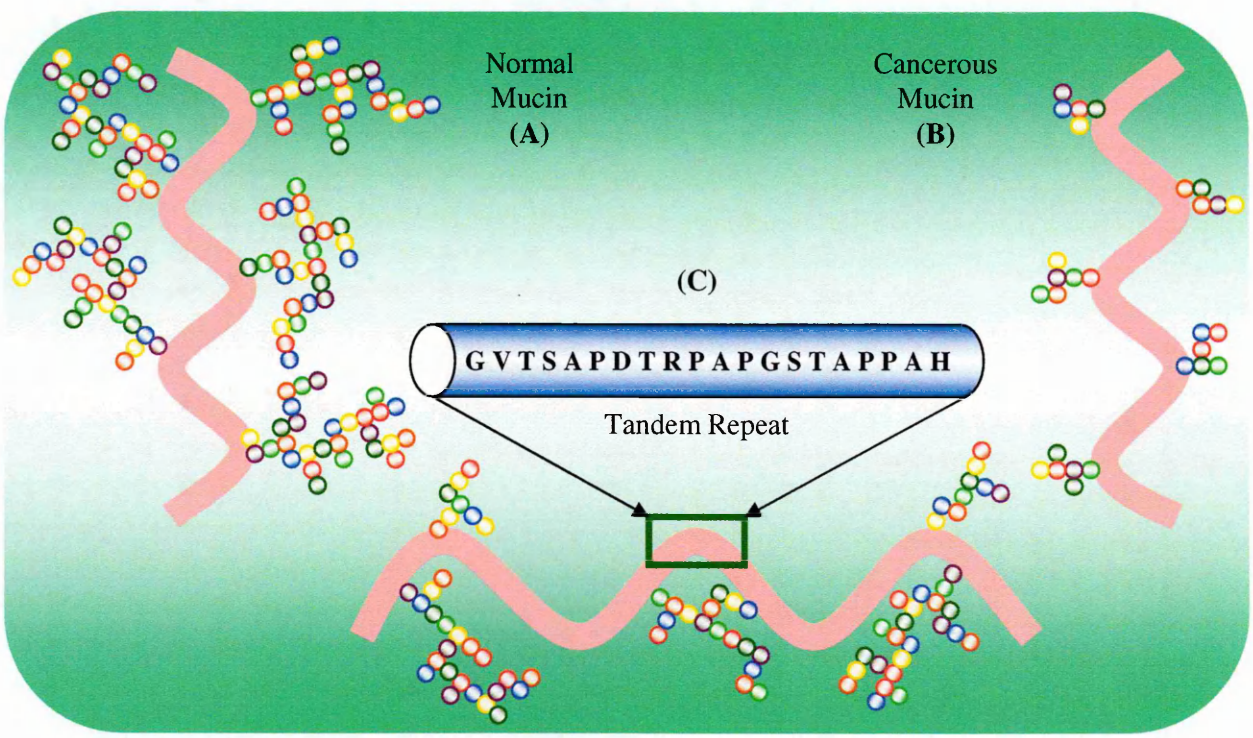


Figure 1.3. Illustration of the glycosylation arrangement on MUC1. (A) Distribution of carbohydrates on normal MUC1, (B) Distribution of carbohydrates on malignant MUC1, (C) The core peptide consisting of a tandem repeat of a 20 amino acid sequence. Adapted from ref [29].

Due to the over expression of MUC1 in carcinoma cells, which also includes the presence of the protein in the same location as adhesion molecules, and the unique length of the extracellular domain extending far beyond the cell surface, MUC1 masks such adhesion molecules in carcinoma cells. This shielding of the adhesion molecules prevents any cell-cell aggregation and cell-matrix interactions from forming. Therefore, in such situations, metastasis is encouraged as the adhesive properties are lost, permitting the separation of the tumour cells from the neighbouring cells and the cell matrix at the primary site. As the tumour cells become unattached from the primary site, the cells are capable of invading the surrounding

tissues, metastasising to form secondary tumours at varying organs around the body. To accompany this, MUC1 may also protect cancer cells from detection and removal of these cells by the body's immune system. This may occur through the combination of the presence of soluble MUC1 from the tumour in the serum, which can inhibit adhesive interactions of migrating cells with endothelial cells, and the ability of MUC1 to shed itself from the transmembrane core. The release of the heterodimer and the soluble MUC1 could reduce the intensity of inflammatory cells at the tumour site, allowing tumour to escape from immune surveillance [29].

Interestingly, MUC1 can also operate as an adhesive molecule, due to the attached carbohydrate structures such as sialyl Lewis^{a} and sialyl Lewis^{x}, in cancers of the pancreas and the colon [24]. These carbohydrates are proposed to be ligands for molecules such as P-selectin. Such interactions are able to imitate the typical mechanism utilised for trafficking leukocytes and in this case to aid in the spread of metastasis. The conventional method of leukocyte trafficking entails a weak binding between the sialyl Lewis carbohydrates present on the leukocytes and the selectins expressed on the surface of epithelial cells. This weak interaction allows a gradual movement of the leukocytes along the cell surface. This motion is mimicked by tumour cells expressing MUC1 containing these carbohydrate structures, also referred to as blood group antigens. Blood group antigens interact with the P-selectin in the same manner allowing migration to take place into other tissues. On the other hand, using the same mechanism, the blood group antigens present on the soluble form are able to bind with the selectin adhesion molecules on the leukocytes, distracting the leukocytes from the tumour. Furthermore, these carbohydrates on the soluble form are able to generate a mucous gel which acts as a barrier and further limits the immune system from gaining access to the malignant tissue [9,24].

Although the general consensus is that greater expression of MUC1 is found on cancerous cells compared to that on normal cells of the same organs or tissue, this does not seem to be the case for all types of cancers. As an example, using monoclonal antibodies (mAb's) against an epitope present on the protein backbone, it was observed that more protein was expressed on colon carcinoma cells compared to normal cells. However, contrary to these findings, when cell lines from both tissues were analysed using

in situ hybridisation, the same levels of MUC1 expression were observed. This strongly suggests that the difference in these results is due to the varying glycosylation patterns on the protein of both types of cells. The carbohydrates present on the normal cells are potentially much greater and longer relative to that on cancer cells, thus masking the mAb from reaching the epitope on the protein, whilst on cancer cells the epitope is more exposed due to the truncated oligosaccharides [16].

The MUC1 protein plays a significant role in malignant cells to evade the immune system and with that advantage to further invade and infiltrate into neighbouring tissues and organs rapidly.

1.4.0 ALTERNATIVE SPLICED VARIANTS OF MUC1

Although MUC1 is one of the main mucins, continual research in the area of mucins has led to the discovery of many MUC1 isoforms generated through alternative splicing, that are often found expressed within the same type of cells or tissues. The different alternative spliced variants vary in structure from the native MUC1 protein in either lacking the transmembrane domain, cytoplasmic domain or the tandem repeat regions. Hence, each of the isoforms also differs in their functional properties. It has also been observed across numerous cell lines that most of the isoforms are overexpressed in cancer cells compared to cells obtained from benign tumours or normal epithelial cells. Whilst the existence of various MUC1 isoforms has been established, there is very little characteristic information relating to any of the known spliced variants available. However, the main reason which contributes towards the lack of information on the functional properties of any isoform has been due to the lack of available antibodies which are capable of reacting specifically and selectively with the individual MUC1 isoforms.

1.4.1 MUC1/SEC

Of the many alternative splice variants, MUC1/SEC is a splice variant which possesses the extracellular domain, including the tandem repeat sequence, of the original MUC1 protein. However, as MUC1/SEC is devoid of the transmembrane domain and the cytoplasmic tail, it is suggested that MUC1/SEC is an original secreted form of MUC1. As MUC1/SEC is secreted directly from the cell surface, it contributes

to the soluble form observed in serum, which also consists of the cleaved product of MUC1. However, this secreted form appears to be copious in comparison to the cleaved product. The secreted form of MUC1 can be differentiated from the cleaved MUC1 by identifying an amino acid sequence that is present at the C-terminal exclusively on MUC1/SEC [20]. Obermair et al [50] demonstrated with primary ovarian cancer cases an interesting expression pattern of MUC1/SEC in benign and malignant ovarian tumour tissue samples. Amongst other MUC1 splice variants, which tested positive, MUC1/SEC has been the single splice variant which displayed negative in all the malignant samples, yet positive in the benign samples though not as intensively expressed as the other splice variants. Data also demonstrated that in ovarian cancers, MUC1/SEC is co-expressed with MUC1/Y in benign tumours, yet in malignant tumours the expression of MUC1/SEC is no longer observed whilst the MUC1/Y remains present. Although the expression of MUC1/SEC is non-existent in ovarian cancers, this does not seem to be the case in cancers of the breast and endometrial carcinoma, where the expression of this secreted form has been clearly detected in malignant tissues [14]. Furthermore, evidence suggests that in breast cancer the MUC1/SEC has the capability to efficiently bind to another splice variant, MUC1/Y, with profound effects on the biological functions [51]. From the investigations performed it seems that expression of the alternative splice variant MUC1/SEC is much dependent upon the type of tissue and the differentiation stage at which the cells are at. Hence, findings from cells or tissue samples from a particular organ are not necessarily observed using a tissue sample from a different organ.

1.4.2 MUC1/Y

Another alternatively spliced variant, MUC1/Y, appears to be expressed significantly in many cell lines and in many cases more than the MUC1. The MUC1/Y protein is identical to the MUC1 with regards to the transmembrane domain and the cytoplasmic tail. However, the splice variant is devoid of the tandem repeat region and the flanking region. The tandem repeat region is a unique characteristic feature of the MUC1 protein and the lack of this feature in the MUC1/Y protein not only has an effect on the overall structure of the protein but also suggests the protein may not portray the typical mucin like characteristics and hence may also possess a different functionality to the MUC1. With the missing tandem repeat domain, the molecular weight of the mature MUC1/Y protein after recycling, appears to be in the range

of 42 – 45kDa [15]. On the other hand, experiments conducted have shown the mature MUC1/Y is able to form a complex with other cell surface molecules, bringing the overall molecular weight to 60kDa. Formation of the MUC1/Y complex shows mediation potentially occurring through the cysteine residues, which are capable of forming disulfide bridges. This splice variant, like MUC1, possesses sequences which are also observed in cytokine receptors, suggesting that MUC1/Y can perform as a cytokine receptor and can possibly be involved in similar ligand interactions. Although MUC1/Y shares similarities with MUC1, the splice variant does not have a heterodimeric component upon where cleavage of the protein can occur [17]. Furthermore, this membrane bound cell protein goes through transphosphorylation within the cell on the tyrosine and serine residues, which are part of the cytoplasmic domain, and in breast malignancy this has been found to occur upon binding to the MUC1/SEC protein. The interaction between MUC1/Y and MUC1/SEC, which occurs at the extracellular domain, has not only thought to induce phosphorylation but has been responsible for instigation of many biological processes such as changes to the cell morphology and prompting of the signalling cascade. Triggers to initiate the signalling cascade are believed to occur from the evidence obtained via the interaction of the cytoplasmic phosphorylated residues to the GRB2 protein, a similar mechanism to that observed with MUC1. One of the many effects of the signal transduction produced through the MUC1/Y protein is the enhancement of tumour progression. Interestingly, in cancers of the breast, the expression of this splice variant is seemingly advantageous as MUC1/Y is apparently found to be present in the malignant tissue, yet undetectable in neighbouring normal tissue of the same organ. However, this is not the case in malignancies of many other organs, where the expression of MUC1/Y is also detected in neighbouring normal tissue of other malignant organs, though in significantly lower levels. In addition, the overall levels of expression of not only MUC1/Y but also MUC1 are likely to show discrepancy in the various breast cancer tissue samples obtained from different patients [15,51].

From the few cell lines studied, the expression levels of the protein MUC1/Y is less variable in comparison to MUC1 in malignant tissues. The presence of MUC1/Y is significantly greater in cancerous tissue and barely detectable in normal tissue, as opposed to MUC1, where the expression ranges from significant to detectable or undetectable levels within tissue samples of the same organ from different

patients. As an example, in observing the expression levels in HeLa cervical carcinoma cells, MUC1/Y was present to a large extent, whereas MUC1 was remotely detectable. It has emerged that in breast cancer tissue and HeLa cervical carcinoma cells to a substantial degree, the expression is limited to only this isoform, suggesting that MUC1/Y could potentially be a promising marker for identifying malignancy. This demonstrates that as MUC1/Y expression levels increase, the level of tumourgenicity also intensifies. However, the opposite is observed with the normal MUC1, lower the levels of MUC1, the greater the association of cells with the malignancy [37]. To further support this evidence Obermair et al [52] have demonstrated the extensive expression of MUC1/Y after screening eight different cervical cancer cell lines, whereby it was found that seven out of eight were expressing this splice variant. Also evident from these studies is that MUC1/Y is frequently expressed with other MUC1 isoforms such as MUC1/X and MUC1/Z.

Similarly, when investigating primary ovarian tumour samples, it was observed that MUC1/Y is much more frequently expressed in malignant than in benign tumours. From the 55 ovarian cancer samples that were analysed, 54 of those samples expressed MUC1/Y, proposing that MUC1/Y is a strong candidate for indicating the presence of malignancy. Furthermore, MUC1/Y splice variant is co-expressed with the splice variant MUC1/SEC in benign tumours. However, this is not the case in malignant tumours [50]. Although this may not be the case with other malignancies, the absence of MUC1/SEC and the presence of MUC1/Y can potentially be a better and more definitive approach in distinguishing the stage of the tumour from benign to malignant for diagnostic purposes and in confirming ovarian malignancy.

The results obtained from the different tissue organs or cell lines explored have shown that the expression of MUC1/Y splice variant is variable in the different cell lines or tissue samples. Thus, although there are many variables with regards to the expression of MUC1 and other splice variants, it is evident that MUC1/Y is associated with many malignancies.

1.4.3 MUC1/X AND MUC1/Z

There is very limited literature discussing the other splice variants, MUC1/X and MUC1/Z, often co-expressed with MUC1/Y. However, though the significance of the role of these splice variants is not exactly known, the structure of these proteins has been established. Similar to MUC1/Y, MUC1/X and MUC1/Z both lack the VNTR region, a distinctive feature of the core MUC1 protein. Furthermore, MUC1/X and MUC1/Z are also found to be recurrent in malignant tumours rather than in benign tumours as in the case of ovarian cancers [50]. Analogous to this is the expression of these short isoforms determined in the cervical carcinoma cell lines, where from the eight cell lines investigated seven expressed MUC1/X and six expressed MUC1/Z. Also from the eight cell lines, the C-3A cell line did not express either of the splice variants including MUC1/Y and MUC1/SEC, yet expression of MUC1 was clearly observed. Furthermore, often found across the cell lines studied, isoforms MUC1/X/Y and Z are expressed in differing combinations together rather than any one of them individually [52].

Between MUC1/X and MUC1/Z, the MUC1/X isoform has been the most examined. Unlike MUC1/Y, the MUC1/X protein does undergo proteolytic cleavage and this takes place in the SEA module within the extracellular domain. The name SEA was derived from the original place where the SEA module was identified in Sperm protein, Enterokinase and in Agrin. Other than the SEA module, which consists of 120 aa, adjacent to the 30 aa N-terminal, the extracellular domain also contains an additional 18 aa, hence overall proving to be a much shorter isoform than the MUC1 protein. From research findings, proteolytic cleavage occurs in MUC1/X at a site that is equivalently present in the tandem repeat region within the MUC1 protein, indicating that the sequence composing the SEA module is the minimum number of amino acids required for proteolytic cleavage to take place [21]. Using cervical HeLa and ovarian carcinoma cells, from preliminary results it has been presumed that MUC1/X may be 2 kDa larger than MUC1/Y, nevertheless further investigations are required to confirm this [37].

Although MUC1/Z is devoid of the tandem repeat, like MUC1/X but unlike MUC1/Y, it has 18 more amino acids within the extracellular region. However, any other structural feature or molecular weight has not yet been established. Schut et al [53] investigated the expression of MUC1 and the splice variants of

MUC1, in particular MUC1/Y and MUC1/Z, on benign prostatic hyperplasia (BPH) and malignant prostate tissue. The findings showed that no significant expression specificity of MUC1/Z was displayed for the malignant tissue over the BPH. However, it has been suggested that BPH could eventually develop into cancer as, like MUC1/Y, MUC1/Z is also preferentially expressed by tumours. In another investigation involving the screening of serous effusions, through the detection of cancer cells in these fluids, it has been found that MUC1/Z along with MUC1/Y are greatly over-expressed in primary breast tumours and ovarian cancer in comparison to the MUC1 protein. Also ascertained is the co-expression of these two splice variants which are rarely detected in benign tissue [54].

It still remains unclear the exact role of the MUC1 splice variants MUC1/X, MUC1/Y and MUC1/Z. However, it has been hypothesised that these considerably shorter forms of MUC1 could have an involvement in signal transduction and possibly other intracellular processes. This could be due to the changes which have occurred in the splice variants within the extracellular domain, whilst the cytoplasmic region has remained identical to the MUC1 protein in all the forms.

1.5.0 THE POTENTIAL OF OLIGONUCLEOTIDES

The vast pressure to rapidly make revolutionary advances within the science and technology field has led to increasing demands in research for the development of novel methodologies in the areas of molecular recognition, diagnostics and therapeutics for a wide range of human disorders. There are numerous novel therapeutic targets which are emerging that may either have a connection with or be the cause of a disease. However, the critical challenge lies in the ability of the ligand or the inhibitor to perform on the basis of selectivity and specificity against the chosen target, to bring about the desired functional result. For this reason, the need for sophisticated pharmaceutical agents has become an essential requirement which has led to the increasing popularity of biologics and oligonucleotide-based applications in particular. Oligonucleotides are molecules which are remarkably multipurpose and have shown to be employed in a broad variety of applications including molecular biology, diagnostics and therapeutics. Oligonucleotide therapies have demonstrated promising outcomes in numerous disorders including cardiovascular diseases (ARC1779 [55,56], NU172 [57,58]) [59], viral (Anti-HIV-1 RT [60] A22 [61]), inflammation

(NX21909 [62,63], Anti-L-selectin [64,65]), angiogenesis (NX1838 [66-68]), neurological disorders (Anti-Substance P [69], Anti-NGF [70]) and many different cancers (AS1411 [71,72], TTA1 [73,74], Anti-PSMA [75]). Oligonucleotides are single strands of DNA or RNA of varying lengths where the ligands are referred to by their own unique terminology. The terminology depends upon their mode of action and application such as aptamers, ribozymes, DNAzymes, siRNA, antisense, and spiegelmers. The DNA / RNA based technologies are still at an early stage with regards to their development and reaching the market in comparison to other established applications like small synthetic molecules, peptides and antibodies. The technologies which fall under the umbrella of the DNA / RNA category are all in progress at various stages with regards to their development. The antisense approach is the most developed and investigated amongst the other methodologies. Currently two oligonucleotide based products are available in the market employing different strategies; antisense and aptamer, to perform the necessary functional action. In 2005, 229 DNA / RNA products had been established by the Datamonitor global research and all these products were active in their research and development programmes. From the 229 products, 19 had demonstrated potential of reaching the market within a decade [76], hence displaying a remarkable potential for oligonucleotide orientated products. The oligonucleotide based technology initially began in 1967 by Belikova et al [77] by putting forward a theory suggesting an oligonucleotide sequence, which is complementary to the mRNA of a target, is able to prevent expression of that mRNA, thus inhibiting the relay of genetic information from the DNA to the protein. Although this was initially proposed for prokaryotes, a decade later Zamecnik and Stephenson [78] in 1978 demonstrated this antisense strategy in a cellular system to block the replication of the virus Rous sarcoma, hence displaying potential for utilising this system in therapeutics [79].

1.5.1 ANTISENSE STRATEGY

The core principle behind the antisense strategy is based on the Watson-Crick hybridisation. A short single stranded oligonucleotide is generated, which is specifically complementary to the mRNA sequence of the selected target. There appears to be a necessity for the oligonucleotide to be of a particular length, based on calculations and studies with different length oligonucleotides. The appropriate length was established to be typically between 15 and 20 nucleotides long [80]. The function of the short antisense

strand is to hybridise with the mRNA target which is responsible for the expression of a particular protein in the cytoplasm. The function of the short antisense strand is to hybridise with the mRNA target which is responsible for the expression of a particular protein in the cytoplasm. This hybridisation is employed to inhibit the translation of the target protein through the initiation of various mechanisms which include; translational arrest of the mRNA when targeted at the 5′ terminus via steric hindrance of the ribosomal activity, impeding the mRNA maturation by preventing splicing or destabilising the pre-mRNA in the nucleus, and degrading the mRNA through the activation of RNase H. The RNase H enzyme is an endonuclease which hydrolyses the RNA strand of the RNA-DNA heteroduplex and degrades the mRNA, whilst the antisense strand remains unaffected [81,82]. In theory, the antisense strategy can be applied to block expression of any protein known to be associated with or responsible for the cause of a disease. Many antisense products have entered the clinical phases and are all at varying stages of development. However, one product which has successfully reached the market is Vitravene, also known as Fomivirsen. Vitravene is an antiviral drug and the first oligonucleotide treatment to be approved by the FDA in 1998 for the treatment of cytomegalovirus retinitis (CMV) [83].

1.5.2 ANTIGENE TECHNOLOGY

An alternative approach is the antigene strategy, which employs a similar approach to the antisense with the essential difference being that the short oligonucleotide strand enters the nucleus to bind with genomic double stranded DNA, forming a triplex helical structure. The triplex forming oligonucleotides operate by sequence specifically recognising the oligopyrimidine•oligopurine present in the major grove of the double stranded DNA. The base pairs involved in Watson-Crick base pairing which form the double stranded DNA helix through hydrogen bonding can still offer hydrogen bond donor and acceptor groups. These groups can specifically engage with the bases of the triplex forming oligonucleotides. This specific recognition leads to the establishment of Hoogsteen or reverse Hoogsteen bonds between the bases pyrimidine or purine of the triplex forming oligonucleotide and the bases within the duplex of the purine strand. The formation of the triplex structure has the potential to inhibit the interaction of numerous protein factors for the transcription phase or hinder the initiation or elongation of the

transcription complex, thus preventing the protein translation process. The distinct advantage of the antigene approach is that only a single target is required for inhibition in each cell in comparison to requiring several hundreds or thousands of mRNA copies as targets in the antisense strategy [84-86].

1.5.3 RIBOZYMES

Another oligonucleotide technology under development for as long as the antisense strategy is the application of ribozymes. Ribozymes are varying lengths of single stranded RNA molecules possessing catalytic activity, making them capable of performing enzymatic cleavage or ligation of other RNA molecules. The ability of ribozymes to selectively catalyse the cleavage of their target demonstrates their potential to be used in specific gene suppression [80,87]. There are six naturally occurring ribozymes that cleave phosphodiester bonds of RNA substrates [88]: hammerhead [89], hairpin [90], hepatitis delta virus (HDV) [91,92], *Neurospora* Varkud satellite [93], bacterial cofactor-dependent *GlmS* [94], and group I intron [95,96]. From these six ribozymes, the two classes of ribozymes; hammerhead and hairpin have advanced into therapeutic development. Hammerhead ribozymes were first discovered in plant viroids and virusoids [97] and have also been found in satellite transcripts of various species [88,98-100]. They are small, self-cleavable, RNA molecules with a secondary structure that possesses a highly conserved catalytic region responsible for connecting three base-paired stems. For the structure of a hammerhead ribozyme to take form, *ca.* 40 nucleotides are necessary, allowing the ribozymes to be fully automated with regards to their catalytic properties and be able to perform *in vitro* in the absence of any cellular organisation [80]. The binding of the hammerhead ribozyme is exclusively based on its sequence, which must specifically complement the sequence of the target RNA, with the structure consisting of three base-paired stems. It is essential that the substrate has the required sequence UH (where H = A, C, or U) adjacent 5′ to the cleavage site. Stems I and III of the hammerhead ribozyme specifically bind with the UH sequence of the substrate to ensure cleavage can take place, whilst stem II is involved in the structural conformation. The self-cleaving catalytic activity of the hammerhead ribozyme ensures complete inactivation of the target's biological function, through the hydrolysis of the RNA substrate [88,101,102]. The second type is the hairpin ribozyme, whose natural catalytic small structure was first found in the negative strand of the tobacco ringspot virus

satellite RNA [103,104]. The hairpin ribozymes are fully self-processing, thus the molecule is able to assemble into an appropriate conformation by itself and perform reversible cleaving of either itself or another RNA strand. For hairpin ribozyme catalytic activity to take place, *ca.* 50 nucleotides are essential which form two stems, each consisting of a helix-loop-helix structure. Thus, the two stems generate four base-paired helices and two loops, with one loop containing the reactive phosphodiester group. One stem is responsible for binding to itself or another RNA molecule and the other stem is present to perform any catalytic reactions. Development of these natural ribozymes can aid in understanding the role of naturally occurring molecules. Also, ribozyme technology can be used to inhibit the replication of many pathogenic viruses and further analyse the functionality of an individual gene in greater depth [105-107].

1.5.4 DNAZYMES

An alternative class of molecules very similar to ribozymes are referred to as DNAzymes. However, in contrast to ribozymes, DNAzymes are not typically found in nature and have greater biological stability as catalytic molecules than ribozymes. DNAzymes, also recognised as deoxyribozymes, are short single DNA strands possessing a catalytic motif and were first discovered in 1994, a decade after the discovery of the ribozymes, by Breaker and Joyce [108]. These molecules are generated through an *in vitro* selection process. A typical structure consists of a central motif with a complementary sequence either side of the catalytic region that is utilised to specifically bind with the mRNA target via Watson-Crick interactions. Once binding takes place, the catalytic motif cleaves the RNA target, thus inhibiting the mutant gene. The catalytic activity is based upon the presence of a metal ion, allowing the cleavage to take place between the purine and the pyrimidine nucleotides [109,110]. The DNAzymes are used to target genes which play an essential role in the development of a disease. Therefore, they are specifically designed to bind to mutant sequences and cleave the relevant target, leaving the wild type unaffected by the entire process. This technology is selected to effect gene suppression and has shown potential in many different disease models, however DNAzymes have not yet advanced in therapeutic applications for dominant genetic disorders [110-113].

1.5.5 SMALL INTERFERING RNA

Small interfering RNA (siRNA) are short double stranded RNA fragments, which mediate the degradation of mRNA, thus inhibiting the gene expression in the RNA interference (RNAi) pathway. This natural RNAi process was first described by Fire et al [114], where it was observed in worms, then also in plants [115], fungi [116], and humans [117]. RNAi involves the presence of double stranded RNA (dsRNA) which is homologous in sequence to the silenced gene and this dsRNA is enzymatically cleaved into smaller fragments referred to as small interfering RNA (siRNA). siRNAs are approximately 19 – 23 nucleotides in length with two nucleotide overhangs at the 3′ end on both sides [118,119]. The 3′ overhang is used to incorporate the siRNA into a protein complex named RNA-induced silencing complex (RISC), which initiates the unwinding of the siRNA into two single strands of RNA; the passenger strand and the guide strand. The guide strand (sense / antisense) will remain associated with the RISC, whilst the passenger strand is degraded. The guide strand will bind with the complementary sequence in the target mRNA molecule and as a result can either induce the degradation of the mRNA target by the activated RISC complex or can block mRNA translation, hence inhibiting protein production [120,121]. In the natural biological systems, the siRNAs are produced in the cell from double stranded RNA precursors by enzymatic nucleolysis to initiate the RNAi process [118,122]. However, this can also be achieved by introducing synthetic siRNA molecules into the living cells of the host and initiating the processes involved in the natural RNAi pathway. This mechanism brings about a specific knock down of the gene expression, resulting in depletion of protein levels responsible for the development or promotion of the disease [117,118,121].

There are numerous technologies available, which are all being developed to investigate and shed light on how different biological mechanisms effect the initiation and progression of life threatening diseases. However, there is tremendous pressure to have a single agent with multiple functional purposes such as detect novel molecular targets, perform as a diagnostic tool and as a therapeutic agent. There is one technology, developed only just over two decades ago, that is making remarkable progress and is showing great potential to meet these multifaceted criteria. With the exception of the numerous

oligonucleotides therapies discussed above, other main therapies available in the market include small organic compounds, peptides and antibodies. To date, antibodies have met the criteria of molecular target recognition, and have been used in diagnostics and therapeutic applications. However, some issues related to antibody technology include the long development time and the lack of complete reproducibility [123,124]. Nevertheless, another novel oligonucleotide based modality, termed aptamers, has entered the drug development market and proficiently made a niche for itself, with currently holding the position of the only class of agents to rival antibodies. In order for aptamers to be considered as valuable molecular tools in all three applications, they must display four main attributes for their selected target. Firstly, the agent must be able to differentiate between the native and mutant form of the target. Hence, the ability to bind specifically to the mutant form leads on to the second characteristic of having the ability to quantify the expression levels of the mutant type. The third trait requires the ligand to perform the main functional purpose of binding with high specificity and affinity in *in vitro* and especially in *in vivo* experiments. Finally, for a ligand to be successful in therapeutic applications, the ligand must be able to inhibit progression of the disease with minimal or no toxic effects. A combination of advances in the combinatorial libraries and high-throughput screening methods have made it possible to accomplish these basic principles for developing suitable drug candidates against many new proteins which are emerging as potential therapeutic targets [3]. Upon establishing ligands which show promising potential, they are further improved through appropriate modifications to improve drug stability and total efficacy under physiological conditions, allowing them to be taken into further comprehensive *in vivo* studies.

1.5.6 APTAMERS

Aptamers can be either RNA, DNA or modified nucleic acid molecules and are short synthetic single stranded oligonucleotides with remarkable properties that can be compared to antibodies. The meaning of aptamers is 'to fit' derived from the Latin word 'aptus' [125] and the Greek word 'haptein' [126] meaning 'to attach to'. Hence, the winning sequences obtained from a vast combinatorial nucleic acid library that are 'attached' to the selected target are designated under the category of aptamers. This class of oligonucleotides portrays many advantages and the one which is of utmost benefit is that they can be generated against a wide array of molecular targets [127]. Targets can range from other nucleic acids [128],

nucleotides [129-131], proteins [66,132,133], peptides [134,135], amino acids [136,137], small molecules [138,139], ions [140,141], cells [142,143] and whole organisms [144]. Also included are proteins which are considered as either unstable, highly toxic or simply too complicated to be purified, especially if they are membrane-bound. Typically the length of the randomised region or of the final binding aptamer sequence, tends to be in the range of 15 – 100 nucleotides [145] long. The final length is sufficient enough to fold via intramolecular interactions including natural base pairing that lead to the formation of secondary and tertiary structures of the aptamer. The secondary structures include short double stranded stems and single stranded loops. An amalgamation of these secondary structures leads to the formation of a reasonably stable three dimensional (3D) structure for binding to the target [146]. The binding mechanism, which is also dependent upon the target, involves a number of different forms of interactions, such as hydrogen bonding, hydrophobic interactions, pi-pi stacking, electrostatic interactions, Van der Waal interactions or shape-shape match, much like the lock and key mechanism [147]. All these interactions contribute towards the aptamers' high affinity and specificity for their target. Aptamers have been able to discriminate between proteins belonging to the same family and differing from each other by only one or a few amino acids. Along with great specificity, aptamers have the potential to have a very high affinity to their target, with an equilibrium dissociation constant (KD) in the nanomolar (nM) or even picomolar (pM) range. These artificial ligands are generated through an in vitro iterative process referred to as a Systematic Evolution of Ligands by EXponential enrichment (SELEX), a methodology to acquire high affinity and specificity aptamers in one complete process.

1.5.7 SPIEGELMERS

Spiegelmers are short single stranded oligonucleotides which are not utilised to target the disease at the gene level, like many of the previous strategies mentioned above. These nucleic acid based molecules are typically raised against targets such as proteins or small biological / chemical molecules which assist in the development, enhancement or spread of an ailment. The spiegelmers approach predominantly branches from the original aptamer technology and are developed using the same methodology employed in the generation of aptamers, with high affinity and specificity for their target. Spiegelmers have come about from the essential requirement to develop therapeutic oligonucleotides displaying great stability, in

particular towards nuclease degradation. These are RNA or DNA ligands in their L – form of chirality, initially introduced in 1996 by Fürste et al, where spiegelmers were synthesised to bind to arginine [148] and adenosine [149]. Nucleotides in their natural state are in the D – chiral form, hence the mirror image of the oligonucleotide displays remarkable stability as they are able to escape recognition by nucleases. However, for the same reason, the selection of spiegelmers cannot be employed using L – oligonucleotides, as enzymatic amplification can not take place. Therefore, this requires an adaptation of the fundamental selection system, using the vital D – enantiomeric form of oligonucleotides during selection. Spiegelmers are thus generated initially by obtaining a mirror image of the target of interest. So in the case of a protein or a peptide which in their natural form are displayed in L -chirality, the spiegelmers are initially selected as the natural D – oligonucleotide against the unnatural D – enantiomer of the target. At the end of the selection process, the winning sequences of oligonucleotides are then synthesised in the L – enantiomeric form, which will then bind with the natural L – target. Thus, this well-designed methodology is able to provide L – enantiomeric oligonucleotides, termed spiegelmers. In addition to their ability to bind to their target with significant affinity and specificity they also demonstrate incomparable resistance to nuclease degradation [150,151].

1.6.0 THE BASIC THEORY BEHIND THE SELEX

The innovative SELEX procedure (Figure 1.4) was developed simultaneously by two independent groups, Turek and Gold [152] and Ellington and Szostak [125] in 1990. The basis of this *in vitro* selection method entails the initial binding of the nucleic acid strands to the target of interest over a set incubation period, followed by partitioning and separation of the unbound molecules from the bound. The bound nucleic acids are then enzymatically amplified, to increase the number of molecules of each species present, which brings a close to the first round of selection. The amplified nucleic acid molecules are then subjected to the target once again for the second round of selection. Typically, the number of rounds of selection are repeated 8 – 15 [153] times to obtain a pool of oligonucleotides which are potentially having high affinity and specificity for the target of interest. Subsequent to the iterative steps of binding, partitioning and amplification, the final pool is cloned and sequenced for the identification of each individual sequence present.

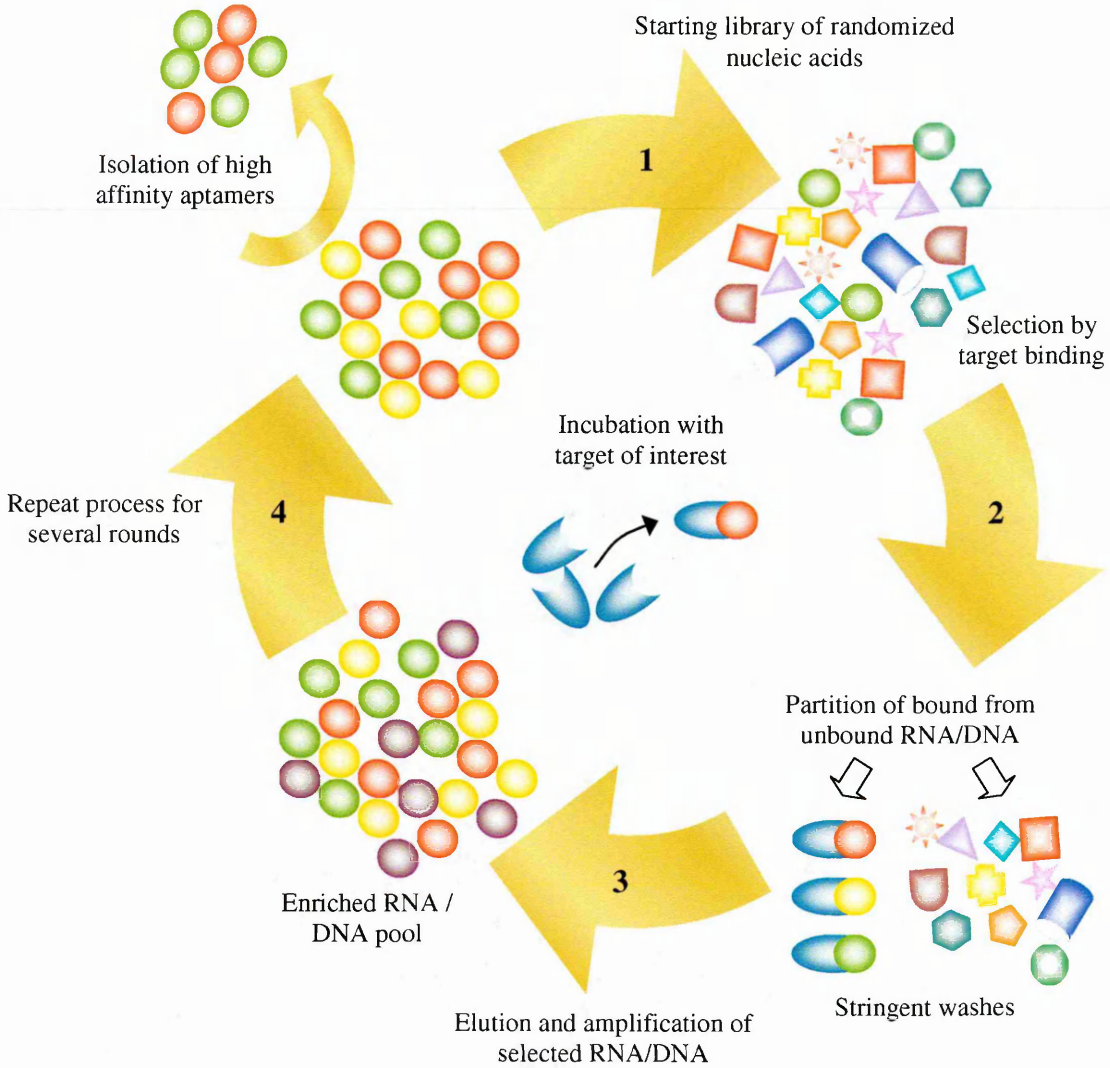


Figure 1.4. An overview of the basic SELEX process. **1)** The starting combinatorial library of nucleic acids is incubated with the target of interest. **2)** The bound nucleic acids are separated from the unbound. **3)** The bound nucleic acids are eluted amplified. **4)** The amplified enriched pool of nucleic acids is subjected to the next round of binding to the target for further isolation. The process is repeated 8 – 15 times. Adapted from ref [154]

The combinatorial library utilised in the SELEX process consists of single stranded oligonucleotides. Each strand has a randomised region of 15 – 75 nucleotides long, with fixed sequences on either side of the randomised region which are necessary for the amplification of each nucleic acid strand. With the SELEX library produced from a combination of all four bases (A, G, C, T or U) and assuming that a randomised region (N) is to be 40 nucleotides long, hence 4^N. Then this would provide a maximum theoretical diversity value of 4^{40} approximately equating to 10^{24} potential different sequences present

within the combinatorial library. However, to simplify the synthesis of the library in any standard laboratory, most researchers are generally limited to employing a library with a diversity of 10^{15} different molecules. Although the library is synthesised to provide a diversity of 10^{15}, taking into account the potential errors during the synthesis that may occur, along with the work up process, the diversity is likely to be reduced down to 10^{14} different sequences. Nonetheless, during the amplification process of the sequences, occasional enzymatic error which may occur leads to an introduction of altered new sequences adding further variety to the original pool of oligonucleotides [146,154].

Once the oligonucleotide sequences have been identified, screened for sequence and structural motif consensus, to reduce cost the sequences are truncated to a minimum length possible without compromising on their affinity and specificity for the selected target. The final length of the aptamer sequences tends to be between 15 – 45 [155] nucleotides long. This length is sufficient enough to form three-dimensional structures such as g-quartets, pseudoknots, stem / bulge and hairpin loops [146], which are considered to play a vital role in binding. In addition to increasing the number of rounds of selection, the stringency of the SELEX protocol can be increased by performing either a negative selection or a counter selection. A negative selection involves performing a round of selection where the library is subjected to the support medium by which the target of interest is attached with. Hence, any aptamers that bind to the support are discarded. The counter selection round entails, incubation of the library against molecules very similar to the target or belonging to the same family. Upon completion of the incubation period, the bound aptamers are discarded and the unbound aptamers are taken further in the following consecutive rounds against the actual target of interest.

The basic SELEX protocol has only been briefly described above. However, aptamers can be generated through a variety of different adapted methods of SELEX. Additionally with the continual advancement of technology, the SELEX method is no longer restricted to just being a bench top process and can now be automated through sophisticated robotic systems. This allows the discovery of aptamers against many different targets to take place simultaneously, a system that dramatically reduces the overall selection time for obtain high affinity and specificity aptamers.

1.7.0 APTAMERS GENERATED AGAINST TARGETS INVOLVED IN CANCER

Aptamers have been generated against a broad range of different types of targets, covering a wide array of diseases and disease promoting physiological processes which include infections, coagulation, inflammation, angiogenesis, and proliferation. However, the following section will focus on aptamers generated against cancer associated targets which have demonstrated to be successful by reaching the pre-clinical and clinical stages of the drug discovery and development process.

The first aptamer to reach the market for the application in human therapeutics and making a new mark by demonstrating outstanding scope in the world of aptamer technology is the anti-$VEGF_{165}$ aptamer. This aptamer is also commonly known as NX-1838, Macugen, Pegaptanib sodium and EYE001 [156]. The aptamer against the vascular endothelial growth factor ($VEGF_{165}$) was initially selected by Ruckman et al [66] and was further developed by Eyetech Pharmaceuticals and Pfizer [157] where Macugen™ is the trade name. VEGF is a growth factor involved in angiogenesis and is responsible for the continuation of healthy vasculature. VEGF is also associated with many diseases including age-related macular degeneration (AMD), heart disease and tumour growth. VEGF has four main isoforms, where the $VEGF_{165}$ is the most important. $VEGF_{165}$ is composed of two domains containing 165 amino acids: a receptor binding domain and a domain involved in the binding to heparin. The anti-$VEGF_{165}$ aptamer is 27 nucleotides long and has been selected to bind to $VEGF_{165}$ through a mechanism where initially the binding occurs with the heparin binding domain (HBD) and then possibly with the receptor binding domain. Macugen shows no significant binding to the other VEGF isoform, $VEGF_{121}$. $VEGF_{121}$ does not contain the HBD, hence clearly suggesting the aptamer specificity for the HBD within the VEGF165, with a high affinity of a 50 pM K_D [158]. The formation of the $VEGF_{165}$-Macugen complex leads to the inhibition of blood vessel growth and vascular leakage, through blocking the $VEGF_{165}$ binding to the receptors on the surface of the endothelial cells. This, in turn, blocks the commencement of an intracellular cascade. The specific inhibitory effect and Macugen's success in clinical trials has led to its FDA approval in the United States in 2004 [156] for the treatment of abnormal growth of blood vessels into the retina (wet – AMD) for retardation in the loss of vision. Although the anti-$VEGF_{165}$ aptamer has entered the market for the treatment of AMD, this aptamer, through the same mechanistic approach, has

displayed potential for tumour growth inhibition. The glycoprotein $VEGF_{165}$ is expressed in normal cells as well as malignant cells. However, the expression of this protein is far greater in cancer cells and is strongly correlated with tumour progression in particular cancers of the breast, prostate, lung, pancreas, stomach, colorectal and melanoma [159]. $VEGF_{165}$ encourages tumour proliferation and metastasis via vasculogenesis, which is the development of new blood vessels in cancer tissues, feeding the total tumour mass and allowing the cells to maintain themselves and perform rapid multiplication. Therefore, blocking of the normal mechanism of $VEGF_{165}$ can lead to the inhibition of progressive cellular proliferation and metastasis. It appears that even a small amount of malignant cell kill, which is involved in the nourishment of the tumour mass, can cause a vital reaction that can halt tumour growth [159]. The effect of the anti-$VEGF_{165}$ aptamer as an anticancer agent has shown to be quite promising. A study performed by Huang et al [160] demonstrated 84% reduction in tumour weight in nude mice cultured with Wilms tumour cells in comparison to the controlled animals after daily injections of the aptamer for 5 weeks. In another human tumour xenograft study involving the monitoring of growth of the A673 rhabdomyosarcoma tumour, after daily injections of the aptamer, on day 16 of the treatment, 74% of tumour growth suppression was observed in comparison to the controlled group [156]. However, the real potential of anti-$VEGF_{165}$ aptamer to perform as an anti cancer agent will depend upon the results which come about during the different stages of the clinical trial phases.

From the range of aptamers selected against oncological targets, the aptamer AS1411, or previously known as AGRO100, has progressed the farthest, by entering Phase II clinical trials. This particular 26 mer aptamer has not been selected through the conventional SELEX method but rather based on observations obtained from previous analyses of guanosine-rich oligonucleotides (GROs). Studies have shown that GROs have the ability to perform anti-proliferative activities in cancer cells and are known to bind to a specific phosphoprotein, nucleolin. This protein possesses a molecular weight of 110 kDa and is mainly found in the nucleolus of proliferating cells. Increasing concentrations of this protein is found in cells, such as cancer cells, which are dividing at rates faster than the normal healthy cells [161]. Amongst its many roles, one of the functional roles of nucleolin is to perform as a vehicle protein, transferring viral and cellular proteins from the cytoplasm to the nucleus or the nucleolus of the cell. This provides an

indication of possibly which destination an anticancer agent will be transported to for effective action to take place. The ability of AS1411 to inhibit cellular proliferative activity comes through the capability of the oligonucleotide to form a stable G-quartet structure. This structure is essential for binding to cell surface nucleolin. Upon binding, internalisation of the aptamer takes place, which is then followed by another intracellular binding with an essential modulator, the nuclear factor – $_K$B (NF_KB). This, in turn, inhibits the activation of NF_KB [72] resulting in the initiation of cell arrest in the S-phase of the cell cycle, causing a halt in the DNA replication process. AS1411 has been able to demonstrate tumour growth inhibition across a wide range of cell lines [162,163] and revealed effective *in vivo* activity in human tumour xenografts [72]. These encouraging results have allowed the aptamer to proceed onto phase II clinical studies for the treatment renal carcinoma and for the treatment of acute myeloid leukaemia, phase II studies are being performed with the aptamer in conjunction with the drug Cytarabine [164].

Another aptamer that is closely following the progress of the nucleolin aptamer AS1411 is the aptamer against the prostate specific membrane antigen (PSMA). PSMA is a transmembrane glycoprotein and over expression of this protein has been well established in prostate cancer cells and in endothelial cells lining the inner surface of blood vessels found in most solid tumours [165,166]. However, PSMA is not detected in normal vascular epithelium. The PSMA expression level has been observed to display a proportional relationship to the aggressiveness of the tumour, as with each progressing stage of cancer the expression of the protein also appears to increase [166,167]. Aptamers against PSMA have been generated by Lupold et al [75] with nanomolar affinity for their target. The aptamers were selected against the extracellular portion of the protein, composed of 706 amino acids. The two selected high binding oligonucleotides are 2′-fluoropyrimidine modified RNA aptamers referred to as xPSM-A9 and xPSM-A10 and they do not share sequence homology. The aim for the aptamers was to inhibit the enzymatic activity of *N*-acetyl-α-linked acid dipeptidase (NAALADase). xPSM-A9 inhibits PSMA in a non-competitive manner with an inhibition constant (K_i) of 2.1 nM and xPSM-A10 inhibits PSMA competitively with K_i of 11.9 nM [75]. From the inhibition studies it has been established that both aptamers bind to separate portions of the extracellular region of the protein and do not compete for the same epitope. Furthermore, results demonstrated that truncation of the xPSM-A9 by more than 5

nucleotides resulted in inactivation of the functional aptamer. However, xPSM-A10 could be truncated by up to 15 nucleotides before loosing its ability to bind to PSMA. The truncated xPSM-A10 retained specificity for the precise epitope present on the target. This was shown by fluorescence microscopy studies, where the aptamer bound to the PSMA expressed by LNCaP prostate cancer cells. In contrast, no binding or uptake of the aptamer was observed on PC-3 prostate cancer cells, which did not express the PSMA protein [75]. In addition to this, the scrambled xPSM-A10 aptamer showed no significant binding to either of the cell lines. Thus, the binding of the aptamer to the target protein is highly sequence specific. The findings from all the experimental studies has allowed the exploration of the aptamer as an *in vivo* targeted drug delivery agent [168,169].

Platelet derived growth factor (PDGF) is another growth factor against which an aptamer has been selected. PDGF is a protein with numerous isoforms and is associated with many proliferative diseases such as restenosis, glomerulonephritis, diabetic retinopathy and cancer [170]. PDGF is a dimeric protein consisting of two homologous chains A and B, which are connected together by three disulphide bridges. The chains present the possibility of generating three different isoforms, PDGF-AA, PDGF-AB, and PDGF-BB. However, in order to activate the biological functions of this protein, the binding of PDGF to two receptors located on cell surface referred to as α and β protein is necessary. α and β both possess five extracellular domains, a transmembrane domain and an intracellular tyrosine kinase domain. The PDGF isoforms containing chain A are capable of binding to receptor dimers consisting of minimum one α receptor. Isoforms containing the chain B are able to bind to dimers with a combination composed of both receptors. Therefore, depending upon the type of binding that takes place between the PDGF isoforms and the type of receptor combinations (αα, αβ, ββ) that are expressed on the target cells, results in the final cellular response that is exerted [132]. Expression of the PDGF isoforms and their corresponding receptors has been observed in numerous cell lines and has been linked to affect many of the cell transformation processes along with tumour growth. This protein has been strongly associated with interstitial hypertension in the majority of solid tumours, which leads to poor prognosis and causes substantial reduction in the uptake of any anticancer agents. The PDGF-B isoform has been found to be responsible for controlling the interstitial fluid pressure (IFP) levels and increased levels can clearly be

detected in solid tumours, whilst the same is not observed in normal connective tissue. This relationship is further supported by evidence demonstrating that the levels of IFP tend to increase proportionally to the tumour size and stage of malignancy [170]. Aptamers have been selected against the PDGF-B to inhibit the protein from binding to the receptor. This should result in reducing tumour hypertension and allowing greater uptake of cancer agents, therefore possibly resulting in tumour reduction. The aptamers obtained at the end of the selection process all share the same secondary structure motif that consists of triple helix with a central conserved single stranded loop region. This loop region is the fundamental motif that is required for high specificity binding and binds successfully to PDGF-AB and PDGF-BB, but has a much lower binding affinity for PDGF-AA. The specificity in this type of binding lies in the presence and detection of the B-chain in either of the isoforms [132]. The propensity of the anti-PDGF-B aptamer to inhibit the binding of PDGF-B to their corresponding receptors was demonstrated in studies performed by Pietras et al [171]. The affinity of the aptamer specifically for PDGF-B is approximately 100 pM and in comparison has a trivial affinity for the PDGF-A, and has shown to reduce IFP in a rat colonic carcinoma model. This result was comparable to that observed using a synthetic small drug inhibitor STI-571. The aptamer was capable of increasing uptake of an anti-cancer agent, Taxol, equally to the drug inhibitor, aiding Taxol in tumour growth suppression [172]. Although the anti PDGF-B aptamer is demonstrating strengths that are equally matching that of synthetic inhibitors, the aptamer is still in the early stages of the preclinical phase of development.

A disulphide linked hexameric glycoprotein found on the extracellular matrix, known as tenascin-C, has been found to be overexpressed in tumour tissue when compared to the expression levels in normal tissue. Tenascin-C is involved in tissue remodelling processes and therefore is present during embryogenesis, tissue repair, inflammation and tumour growth [173]. Aptamers were developed by Hicke et al [74], from which one aptamer referred to as TTA1 was reduced in length to 39 nucleotides with a final molecular weight of 13.4 kDa from the parent aptamer sequence of 71 nucleotides. TTA1 has a high affinity for tenascin-C with a K_D of 5 nM, whereby the binding of the aptamer to the target occurs through structural specificity. The folding of the aptamer adopts a secondary structure with the arrangement of three stems (I, II, III) branching from a central junction, which is considered to be the

essential binding region. It appears that the specific sequence involved in the make up of stem I is not crucially important as is not the closing loops formed at the end of stem II and III. However, stems II and III are vital to allow recognition and binding to take place to the tenascin-C protein. In addition, stem I plays an imperative role with regards to the secondary structure. Stem I contributes towards the overall structural stability of the entire aptamer and sustains the binding potential presented by stem II and III regions. Studies have shown removal of stem I has led to loss of binding. However, if stem I is maintained with a different sequence then the binding remains unaffected, therefore clearly suggesting the structural significance of stem I possess no sequence specific relevance [174]. An *in vivo* investigation for the delivery of radioisotopes or cytotoxic agents by TTA1 has shown to be promising, as there was aptamer uptake by a variety of solid tumours including breast, glioblastoma, lung and colon. The uptake of the aptamer by the tumour, accompanied by the rapid clearance from the blood and the other non-target bearing tissues allowed clear tumour imaging to take place [175]. The development of TTA1 is still near the beginning and has many *in vitro* and *in vivo* experimental studies to go through and demonstrate successful results in these investigations before the aptamer can enter clinical trials.

The status of any therapeutic agent that has entered the clinical trials can be monitored through the website www.clinicaltrials.gov. Table 1.0 summarises the aptamers developed mentioned above against their targets along with their development status.

Name of Aptamer	Target	Type of Oligonucleotide	Clinical Phase
Macugen (NX1838, Pegaptanib sodium, EYE001)	$VEGF_{165}$	RNA	Phase IV
AS1411, AGRO100	Nucleolin	DNA	Phase II
xPSM-A10	PSMA	RNA	Preclinical
Anti-PDGF-B aptamer	PDGF-B	DNA	Preclinical
TTA1	Tenascin-C	RNA	Preclinical

Table 1.0. Aptamers present in clinical and preclinical phases. The targets to which these aptamers have been generated against have all shown strong association with distinctive malignancies.

As seen from these few examples, aptamers have made remarkable progress since their discovery in the early 1990's. There are many other aptamers that have been and are currently being developed against numerous different targets associated with various carcinomas which are in early development. However, undoubtedly, there is great scope for aptamers to follow suit of the ones that have already proven to be a success in the world of diagnostics and therapeutics.

1.8.0 APTAMER ADVANTAGES Vs ANTIBODIES

Up till the mid 1990s small molecules have been the most dominant in the area of therapeutics, whilst antibodies have been leading in the area of medical diagnostics and analytical applications. However, the discovery of aptamers has taken the approach to diagnostics and therapeutics to a new level, as they are reasonably cost effective and much of their synthetic processes are automated. Aptamers have been directly compared to antibodies with respect to their functional attributes. The overall ease in application of aptamer has given them many positive characteristics. One of the greatest advantages of aptamers is that the entire selection process does not require the use of animals. Hence, this saves time, as typical antibody production takes up to six months from the point of animal immunization. Other benefits include significant reduction in costs, as the expense of animals and their maintenance is eliminated. Furthermore, an *in vitro* selection method allows generation of aptamers against practically any target. This also includes targets which may not bring about an immunogenic response in animals, due to the target possessing a similar structure to that of the endogenous molecule. Toxins can also be used as a target, which can be problematic in an *in vivo* system, since the toxins can cause animal death [153]. Aptamers can easily be site specifically labelled with a fluorophore or modified for pharmacokinetic purposes, and modifications can be performed in pH ranging typically from 4 to 8.5 [155], temperature or solvents necessary, as aptamers are stable in non-physiological conditions. Antibodies on the other hand are sensitive molecules, and cannot be subjected to non-physiological conditions. Hence, labelling a particular amino acid on antibodies can prove to be a difficult task [176]. Monoclonal antibodies (mAbs) are very large molecules and are restricted to extracellular targets. Hence, mAbs take a considerable amount of time to penetrate through the tissues and are not proficient in an intracellular environment, as mAbs aggregate and misfold in the cytoplasm [170]. Antibodies also possess prolonged blood residence, taking up

to several days to clear from the blood. This limits the effectiveness of the short-lived radionucleotides and therefore results in poor *in vivo* image quality [177]. Table 1.1 outlines more advantages and disadvantages of aptamers relative to antibodies in all different aspects from synthesis and modification to applications.

1.9.0 CHEMICAL MODIFICATIONS OF APTAMERS

Aptamers have shown remarkable versatility in many of the in vitro applications. However, their drawbacks emerge when they are required as therapeutics agents in animal and human models. Other than having a high affinity and specificity for their target, aptamers must demonstrate stability against nucleases in the blood. Also, aptamers must be able to remain in blood circulation for a period of time sufficient enough to bring about a therapeutic dose response before entering renal clearance. Prior to any modifications, DNA and RNA strands are highly unstable in plasma, as nucleic acids are subjected to exonuclease and endonuclease degradation. Aptamer stability can be measured in vitro and studies have shown a typical RNA oligonucleotide to possess a half-life of only a few seconds, whilst DNA strands can remain stable from 30 minutes to an hour [154]. Modifications to nucleotides can dramatically improve plasma stability and these modifications can be pre-SELEX or post-SELEX. In the case of RNA oligonucleotides, the library will typically consist of modified RNA strands, as RNA molecules are naturally more unstable than DNA molecules. As RNAase found in blood specifically commences degradation from the pyrimidine ribonucleotides, the most common modification incorporated in the RNA library is the substitution of the 2' hydroxyl group (2'OH) on the ribose with either a 2'Fluro (2' F) or a 2' amino (2' NH2) group.

Pieken et al [178] have demonstrated effectively that modification of the sugars on the pyrimidines have significantly increased resistance towards nuclease degradation for a prolonged period of time. Another approach to prevent the action of exonucleases for RNA and DNA is to cap the strands at the 5' and 3' ends. Through the capping of the 5' end of the oligonucleotide, exonuclease activity in the direction of 5' → 3' is limited and similarly by capping the 3' end, degradation by the 3'→ 5' directed exonucleases are

inhibited to certain degree. The most common modification for the 3' end is the capping with an inverted thymidine (3'-dT) [65] and biotin at the 3' terminus can also be attached [179]. Having only a stable aptamer is not sufficient to meet the demands of good pharmacokinetic properties.

Aptamers	Antibodies
• Aptamer selection is a chemical *in vitro* process	• Antibody production employs a biological system, requiring the use of animals
• Time taken to complete the selection process varies between days to few weeks	• monoclonal antibody generation takes 6 months from the time of animal immunization
• *In vitro* process allows selection to take place against any target proteins, including toxins	• Problematic if a protein target is similar to an endogenous protein. Hence, low immunogenic response. Toxins can result in animal death
• Selection can be performed under any conditions, i.e. varying temperatures and pH. Ideal for *in vitro* diagnostic assays	• Can only be formed under physiological conditions of the animal
• Target site of the protein is determined by the investigator	• Target site of the protein is determined by the immune system
• Has minimal or no batch to batch variation	• Variation of activity observed from batch to batch
• Diverse selection of chemical modifications can be performed for the desired functional purpose	• Limited possibilities for modifications
• high binding affinities in the nM to pM range	• high binding affinities in the nM to pM range
• Great specificity, able to distinguish with a difference of only a few amino acids	• Polyclonal antibodies are not highly specific for a epitopes on the target
• Small size with good tissue penetration. Rapid blood clearance providing higher signal / noise ratio in immunoassays	• Large molecules with slow uptake. Blood clearance can take up to several days to obtain a significant signal / noise ratio
• Small in size, aptamers can be selected for extracellular and intracellular targets	• Large in size, selection restricted to extracellular targets as penetration issue prevents accessing intracellular targets
• Stable over a broad range of pH, temperature and solvents. Regenerated upon denaturation	• Temperature, pH and solvent sensitive. Irreversible denaturation
• Long shelf-life. Can be transported at ambient temperatures	• Limited shelf life. Needs special care in storage. Transported only lyophilised / on ice
• Non-toxic and lack immunogenicity	• Non-toxic and high immunogenicity
• Excellent solubility. Injected in volumes as little as 500µl	• Poor solubility. Requires large volumes for dosing
• Not readily available for any given target. Generation depends on properties of the target. Significant quantity of the target is required for aptamer development	• Not readily available for any given target. Generation depends on properties of the target. Significant quantity of the target is required for antibody development
• No one ideal method is available for aptamer development. Conditions need to be varied depending upon the target	• Same generic method is employed for antibody production of using physiological conditions

Table 1.1. Properties of aptamer in comparison to antibodies. A large combination of positive attributes placing aptamers as rivalry candidates of antibodies. Adapted from ref [124,180,181].

It is vital that the aptamers remain in circulation and are not eliminated through the kidneys within minutes. The length of aptamers selected for therapeutics tends to be in the range of 25 – 40 nucleotides long. This usually equates to a molecular weight range of 8000 – 12000 daltons (Da), where this size has demonstrated to be far too small to display a reasonable rate of clearance in animal models. The problem of rapid renal clearance is tackled through further modifications on the aptamer, which involve the attachment of a large amphiphilic molecule such as polyethylene glycol (PEG). Various different molecular weights of PEG are available and can be linked to the aptamer via several methods, the most common of which is the attachment through a primary amine on the 5' terminus. The attachment of PEG on the 5' end of the oligonucleotide combats two objectives simultaneously; firstly linking of the PEG caps the 5' terminus to prevent attack from exonucleases and secondly this conjugation holds the potential to increase circulation of the aptamer from hours to days [182]. Another strategy that is recently becoming relatively popular is the utilisation of liposomes. In some cases, a combination of liposomes and PEG has been employed where the liposome plays the role of a drug delivery vehicle. Modifications for the intentions of other than stabilisation and improved pharmacokinetic properties include attaching a fluorophore for signalling or fluorescence based studies and the linking of biotin for immobilisation purposes. Figure 1.5 indicates the possible modifications which can be performed on aptamers to best suit the functional requirement either as a molecular, diagnostic or a therapeutic agent.

1.9.1 MODIFICATION OF NUCLEOTIDES

Nucleotides are the core of the nucleic acid structure, consisting of three fundamental components; a base (purine / pyrimidine), a sugar moiety and a phosphate group, which are all linked together. These three individual components provide a platform where modification to particular moieties can take place and many modifications can be incorporated during the chemical synthesis of the nucleic acid strand. Chemical dressing of aptamers has vastly developed in the recent years, especially for RNA aptamers, due to their severe lack of stability in most non-sterile conditions and biological fluids.

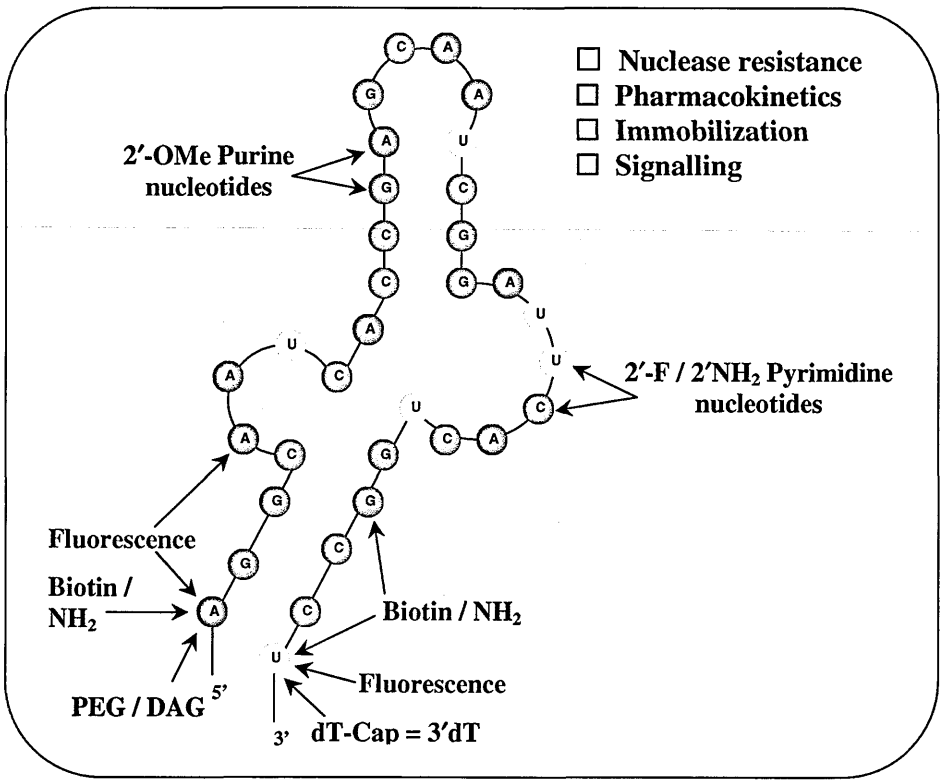


Figure 1.5. Overview of potential aptamer modifications. The modifications outlined provide solutions for nuclease degradation, improving pharmacokinetics, immobilisation to solid supports and signalling for fluorescence studies. Adapted from ref [179].

There are two approaches of modifications which can take place. One where modifications are already incorporated into the main oligonucleotide library prior to selection or the other where the necessary or desired modifications are carried out on only the winning aptamer sequences. The first approach is principally more appropriate for RNA aptamers, as it is absolutely necessary that the RNA library remains fully intact throughout the entire selection process. Only few modifications can be used in this approach, as the majority of these modifications is on the backbone of the nucleic acid and can interfere with enzymatic recognition involved in the amplification step of the SELEX process. Hence, the most popular modification is the exchange of the 2′OH on the pyrimidine ribose sugars, which is involved in a nucleophilic attack on the adjacent 3′ phosphodiester bond, with a small group that is either a 2′F or 2′NH_2. These 2′ ribose sugar modifications are adequate to confer the RNA strands nuclease resistance. However, to further enhance the stability of the selected RNA aptamers, modifications on the purines can be included during their chemical synthesis. The modification on the purine nucleosides involving the

substitution of the 2′OH on the ribose sugar with the O-methyl group (O-Me) provides protection against less significant RNAases present in the blood [183]. Typically, half of the 2′OH groups on the purine ribose sugar can be substituted and in rare circumstances all the 2′OH groups can be substituted with O-Me groups. This substitution does not affect the original binding properties of the RNA sequences. Ruckman et al [66] demonstrated this modification by substituting all the 2′OH groups on the purine ribose sugars with 2′OMe groups on the VEGF aptamer, which already had 2′F pyrimidine modifications prior to SELEX. Another type of modification, though not as widely used as the 2′F or 2′NH_2 is the phosphorothioate modification. This involves the substitution of the oxygen present in the phosphate backbone, the oxygen that is not involved in the bridging of the sugars, with a sulphur atom. Like the phosphodiester, the phosphorothioates also provide a negative charge to the backbone, but provides greater protection against nuclease degradation. The negative aspect of phosphorothioate modification is that this modification leads to non-specific binding to proteins, along with the formation of a phosphorous chiral centre, resulting in numerous enantiomeric forms, hence counteracting the positive aptamer attributes [184,185]. Another alternative modification is similar to that of the phosphorothioates. In this case the same oxygen on the phosphate backbone is substituted with a methyl group and this form of strand is referred to as methylphosphonates [186,187]. The methylphosphonates have the same advantages and disadvantages with the phosphorothioates in providing resistance against nucleases and chiral formation, with an additional drawback. The solubility of the entire molecule is lowered due to the lack of charge on the phosphate. A further implication of this modification is the occurrence of steric hindrance, caused by the presence of the methyl groups, which leads to inadequate duplex formation [185]. Other than the 2′ position of the sugar, the C_5 position of pyrimidines and C_8 position of purines has also been investigated for modification. Position C_5 of pyrimidines have had bulky groups, such as benzoyl groups, attached to the base which have shown no evidence of interfering with the enzymes employed in the SELEX process. These forms of more elaborate chemical dressing of the bases have been summarised by Bruce Eaton [188] and Booth et al [189]. The basic and most common chemical modifications have been illustrated in figure 1.6 within a short oligonucleotide sequence of 5′-TACG-3′.

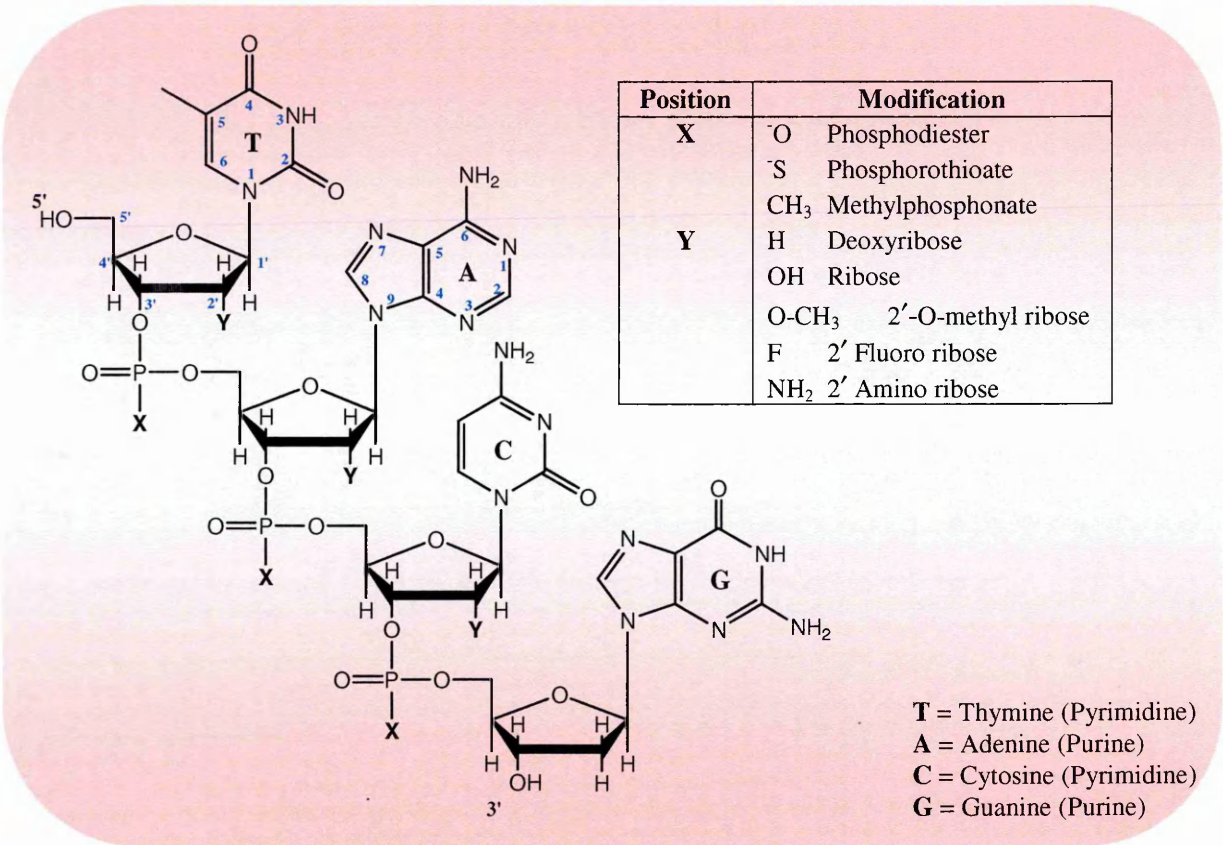


Figure 1.6. Possible sites for common chemical modifications. Structure of a short oligonucleotide sequence of 5′-TACG-3′ indicating the main substitution sites: **X** = Backbone modifications and **Y** = sugar modifications. The pyrimidines can be modified at position C_5 and the purines at C_8. Adapted from ref [185].

Most DNA aptamers are stable prior to any modifications in a variety of conditions and have demonstrated relatively good stability even in biological fluids for a limited period of time, allowing DNA aptamers to be used directly for many *in vitro* studies. Nonetheless, for applications in diagnostics or *in vivo* work, modifications such as terminal capping using an inverted thymidine residue, aiding in the protection against exonucleases, would be inevitable. Attaching functional moieties at the end of the DNA strand, such as an amine linker, allows conjugation to fluorophores (Cy3, Cy5, TAMRA, FITC) hydrophobic molecules (PEG) [179], providing a means of attaching the aptamer to a solid support and offers protection towards exonucleases. Following aptamer selection, modifications to DNA aptamers can be directly incorporated during the chemical synthesis. The critical challenge remains in the

attachment of large hydrophobic molecules for which currently an external ordering service and automated synthesis is not available.

1.9.2 PEGYLATION

In addition to overcoming the drawback of nuclease degradation in aptamers, the second predicament to resolve is to reduce rapid renal clearance and the obvious method has been through the attachment of a large hydrophobic molecule. The most common and versatile macromolecule employed for such a purpose is polyethylene glycol (PEG) and through a process referred to as pegylation it allows the covalent linking of PEG to the aptamer through variety of different chemical reactions. The pegylation process was first developed by Abuchowski et al [190] in the late 1970s using low molecular weight 1.9 kDa and 5 kDa PEG molecules to covalently attach to a protein via the amino groups. Abuchowski and his group observed enhanced circulation of the pegylated protein in the blood and displayed no indication of any onset to an immune response. Furthermore, PEG did not display any adverse effects with reference to tissue or organ damage. The polymer has numerous attractive properties such as being non toxic and non immunogenic. Furthermore, the polymer is neutral and does not hold an overall charge, permitting PEG to be amphiphilic by nature. This renders the polymer soluble in aqueous solutions as well as in organic solvents. Lastly, PEG also possesses a chemically inert backbone which does not react during chemical activation of the PEG terminus.

Pegylation presents several advantages, and the main advantage is the increase in molecular weight of the therapeutic molecule, which significantly improves the blood circulation time and *in vivo* half life of the pegylated molecule and reduces rapid renal clearance. The prolonged in blood circulation suggests a far slower absorption rate into the tissues, where saturation of the receptor will also be at the same rate. Thus, the pegylated therapeutics cannot quickly and easily diffuse back out of the tissues into the blood circulation for excretion. Although pegylation encourages greater blood circulation of the therapeutic drug, the modified molecules are not found to be accumulated in large amounts in the reticuloendothelial system (RES) organs, now more commonly known as the mononuclear phagocyte system, the liver or the spleen. A constructive knock on effect of the enhanced circulatory half life is the increase in potency,

which clearly improves patience convenience as the frequency of dosing can be reduced, providing greater patience flexibility. Shielding and protection of the therapeutic drug is another direct benefit of pegylation. For example, the large polymer will mask the therapeutic surface from degrading enzymes and protect the drug like molecule from any other interfering substances by means of steric hindrance [191,192] (figure 1.7). The excretion of the pegylated molecules is highly dependant upon the total molecular weight and PEG is cleared *in vivo* without any structural change. It has been found that a molecular weight below 20 kDa is typically cleared in the urine, whilst larger PEG moieties in excess of 20 kDa are removed at a slower rate through the urine and the faeces. Evidence indicates that small PEG oligomers of molecular weights of less than 400 Da are susceptible to *in vivo* degradation by alcohol dehydrogenase, to metabolites which are toxic. However, conversely, the toxicity of PEG has in essence been diminished with oligomers above the molecular weight of 1000 Da. This has further been reinforced by the examples of the FDA approval of the use of PEG in pharmaceuticals, foods and cosmetics[193].

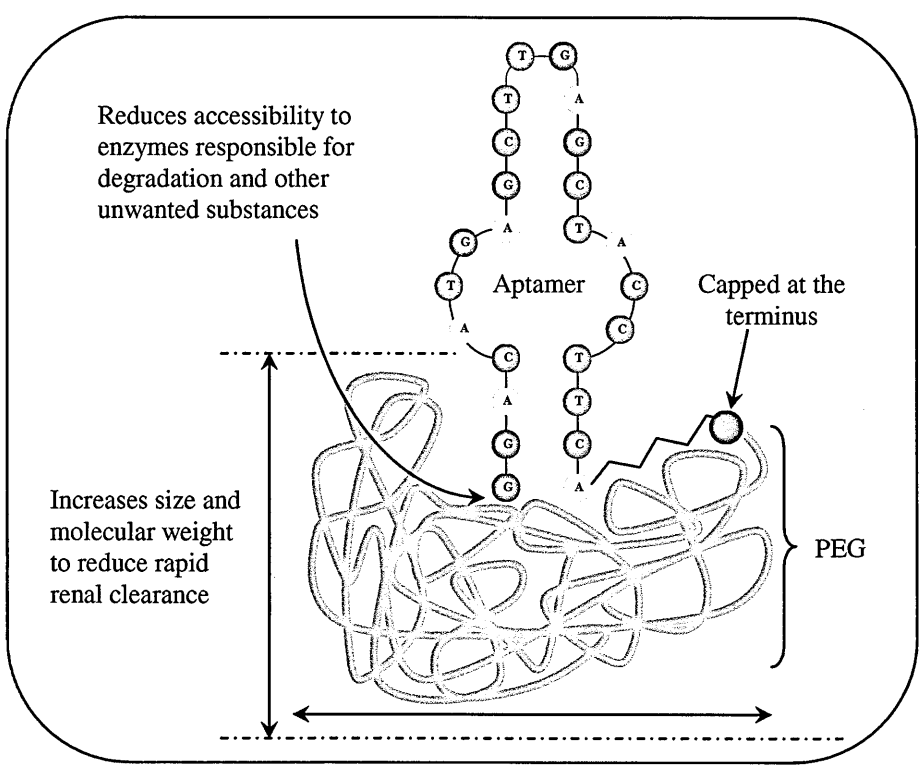


Figure 1.7. The key advantages of pegylated aptamer. An illustration of PEG able to shield the vulnerable termini of the aptamer from degradation initiation most importantly increases the size of the overall therapeutic drug to limit rapid renal clearance. Adapted from [192].

Although there are enormous advantages of conjugating PEG to the desired therapeutic molecule, there are a few minor drawbacks with the use of PEG, which need to be taken into consideration. Firstly, as the PEG polymer is chemically synthesised by the means of linking ethylene oxide subunits together, there is strong likelihood of PEG polymer batch to batch variation occurring as each batch will have different number of ethylene oxide subunits [194]. This is referred to as polydispersity. Polydispersity will affect the pool of pegylated drugs, consisting of varying molecular weights that can lead to assorted biological properties with respect to *in vivo* circulation times. Nonetheless, the extent of polydispersivity is not as large due to recent improved synthesis and purifications methods. However, this limitation needs to be taken into consideration if low molecular weight drugs are to be conjugated to PEG, as the overall molecular weight of the molecule plays a crucial role in its required therapeutic effect. The second issue is related to the excretion of the polymer, as typically PEG is cleared from the body either through the urine or faeces. However, very high molecular weight polymers, though exact weight threshold is unknown, can lead to accumulation in the liver causing macromolecular syndrome. The kidney clearance threshold for proteins is approximately 60 kDa, but the exact kidney excretion threshold for PEG cannot be determined through an extrapolation from different molecular weight proteins, as this is dependent upon a single factor. This important factor is the total molecular weight of the polymer after all the water coordination, which drastically increases the volume of the polymer by 3 - 5 times that of a protein of the same molecular weight [192]. This reduces the kidney clearance maximum for polymers and to some extent restricts PEG's normal movement across the glomerular membrane through a 'snake like' movement. Another potential drawback is the way the therapeutic drug is dressed with PEG. For example, PEG can either be conjugated to the drug at single site or at multiple sites and both options have different consequences. PEG attached at a single site is most likely to retain the original activity of the active site if attachment is carefully chosen, as the effect of steric hindrance and the possibility of the attachment occurring at the binding site is reduced. On the other hand if pegylation is to occur at multiple sites on the therapeutic molecule, then the chances of steric hindrance and attachment at the binding site is increased resulting in a probable reduction or a total loss of bioactivity [192,194].

Each ethylene oxide monomer, which is used to generate a larger PEG polymer, is represented by a whole digit *n* which has a molecular weight of 44 Da. Therefore the result of *n* multiplied by the molecular weight of ethylene oxide (*n* x 44 Da) would equate to the average molecular weight of the full PEG chain length. The general structure of the polymer is: $HO - (CH_2CH_2O)_n - CH_2CH_2 - OH$, with both ends terminated with hydroxyl groups and from this structure the PEG polymer can either be linear or branched (figure 1.8). A common modification which is often carried out on the general structure of the polymer for the purpose of simplifying the functionalisation of PEG is converting one of the hydroxyl groups to a methoxy group, generating monomethoxy PEG (mPEG). mPEG has a typical structure of $CH_3O - (CH_2CH_2O)_n - CH_2CH_2 - OH$.

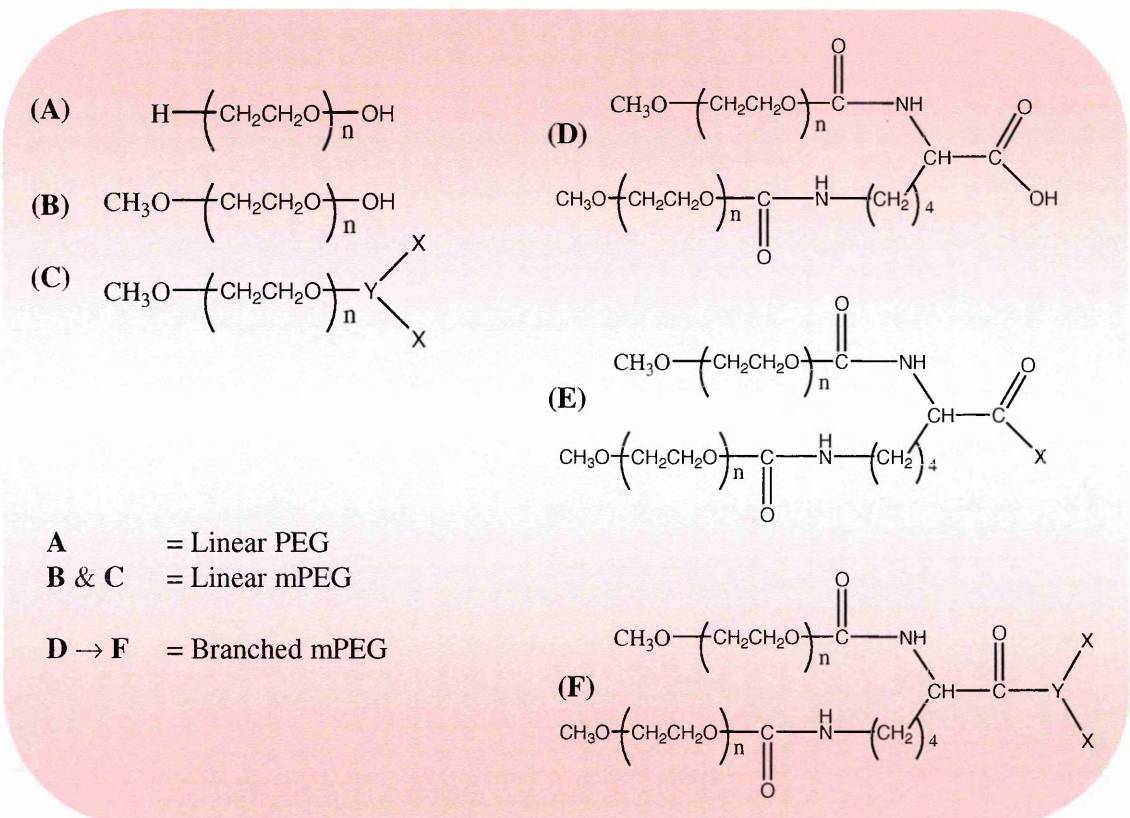


Figure 1.8. Various linear and branched structures of PEG. (**A**) Linear PEG, (**B**) linear mPEG, (**C**) functionalised linear forked mPEG, (**D**) branched mPEG, (**E**) functionalised branched mPEG, (**F**) functionalised branched forked mPEG. **Y** = a group carbon branch moiety and **X** = a linker or a functional group for conjugation. Adapted from [193,194].

This creates a unifunctional polymer with a functional group only on one end, eliminating the potential of cross-linking from occurring during conjugation to the therapeutic molecule, hence avoiding formation of unwanted by-products. The polydispersity (MW/M_n) of PEG is fairly low, as for PEG chains with a molecular weight of less than 5 kDa is ~ 1.01, and for PEG polymers with a molecular weight of greater than 50 kDa the polydispersity is ~ 1.1. Independently of the type of PEG chains, linear or branched, each ethylene oxide unit present within the PEG chain is capable of binding to 2 – 3 water molecules. This increases the overall volume of the polymer and further aids in the action of shielding and masking. However, it appears that the branched chain of PEG has a greater gain than the linear PEG, as the branched PEG along with increasing the molecular weight can also increase the overall size of the molecule without the need to pegylate the therapeutic drug at multiple sites. Furthermore, as the branched PEG comes across larger in size than the linear PEG of the same molecular weight, it provides the possibility to employ a smaller molecular weight branched PEG than if it was a linear PEG chain [193,195].

An mPEG or PEG with hydroxyl groups at both termini does not allow conjugation to take place between the polymer and the target molecule directly. It is necessary that the hydroxyl groups either at both ends or at one terminus are first activated with a reactive functional group. Functionalisation of the native PEG is performed through a relevant chemical modification of the hydroxyl group. The type of chemical modification carried out is dependent upon the functional group that is selected on the basis of the reactive group present on the therapeutic molecule. There are numerous chemical modifications which can be carried out to activate the PEG terminus with electrohphilic functional groups such as active ester, active carbonate, aldehyde or tresylate (figure 1.9). These activated functional groups of PEG are able to react with specific groups of the therapeutic molecule such as amines, sulphydryl group or other nucleophiles for conjugation [195,196].

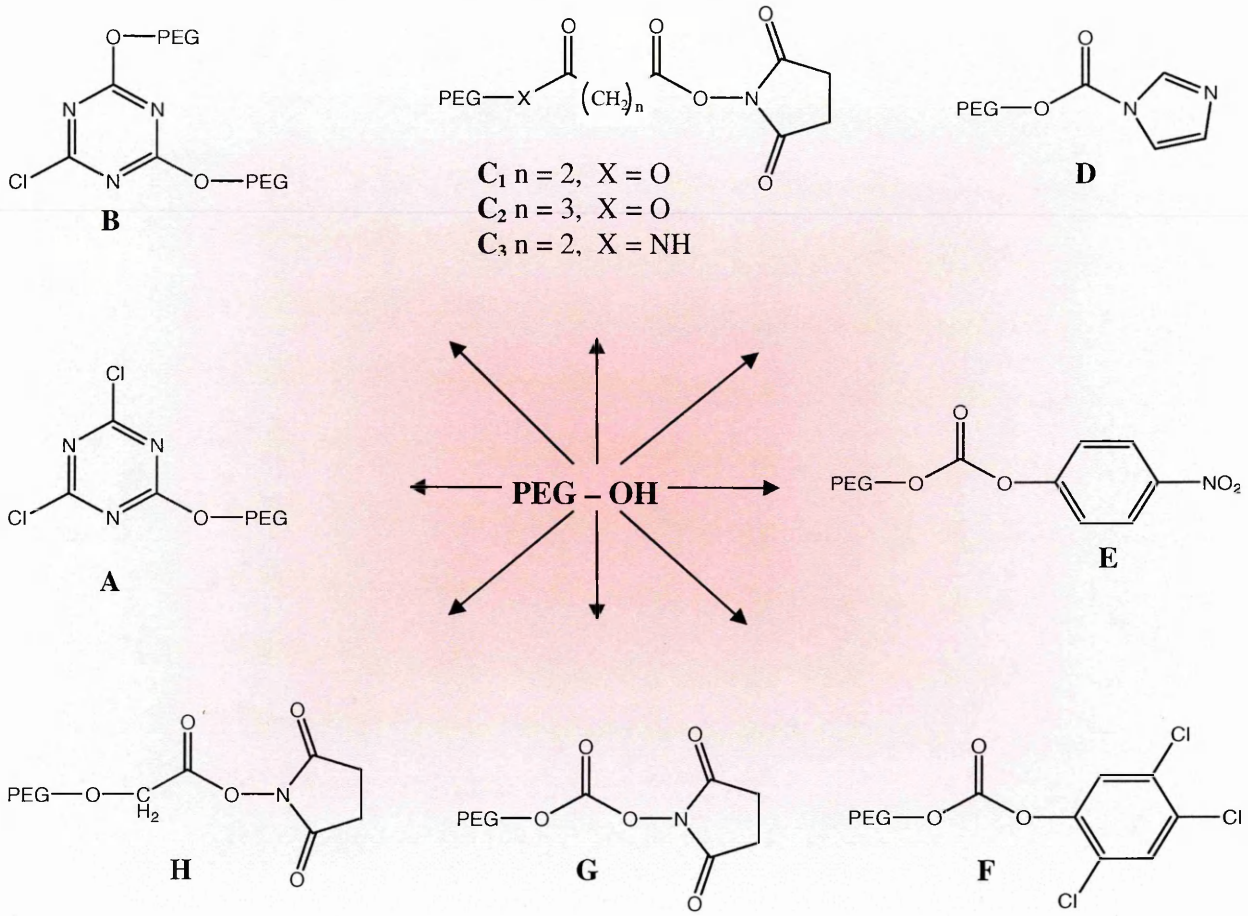


Figure 1.9. Functional groups for activated PEG. **A**) Cyanuric Chloride, **B**) variation of cyanuric chloride, C_1) PEG-succinimidyl succinate, C_2) substitution of succinate by glutarate, C_3) substitution of aliphatic ester in C_1 by an amide bond, **D**) imidazoyl formate, **E**) and **F**) variations using phenyl carbonates of PEG, **G**) succinimidyl carbonates of PEG, **H**) succinimidyl active ester of PEG. Adapted from [196].

1.9.3 PROJECT AIMS

The isoform MUC1/Y has been selected as a tumour biomarker for the development of high affinity and specificity aptamers. This isoform has been the most investigated amongst many known MUC1 splice variants. Although aptamers have been developed against the MUC1 protein [197], it is expected that aptamers against MUC1/Y can be far more promising and advantageous for a number of reasons.

The make up of the MUC1/Y splice variant is far less complicated compared to the MUC1 protein, due to the lack of the tandem repeat region, which results in the absence of the carbohydrate structures that

are characteristically present. This has resulted in the loss of adhesive properties typically present in the MUC1, allowing over-expression of MUC1/Y to take place with the ability to invade far more easily into the lymph vessels and metastasise to the lymph nodes and other organs, proving to be a more aggressive isoform of MUC1. Furthermore, the lack of the tandem repeat region eliminates the possibility of false negative results caused by the inconsistent glycosylation pattern and therefore resulting in a variation in the exposure of the specific epitope expressed on the MUC1 protein.

An additional structural advantage of MUC1/Y is that this isoform does not undergo proteolytic cleavage, hence the protein is not shed from the originating cellular surface. For this reason the interacting therapeutic agent, which in this case will be an aptamer, has a greater probability of reaching the tumour site. By eliminating the effect of having futile aptamer circulating in the bodily fluids, this could provide a better scope for a greater tumour uptake of the aptamer and increases its potential effects as a therapeutic agent.

The main aim of the project is to generate aptamers against the MUC1/Y splice variant, with a dual motive. The first aim has been to develop a selection method which is a practical bench top process, simple with respect to the facilities required and can be completed within a few days as opposed to several weeks. The basic SELEX method, which on average consists of 10 rounds of selection, can take as long as twenty days to generate the final pool of potentially high binding aptamers. However, the key objective here would be to develop a selection protocol which generates high binding and high affinity aptamers, and the entire selection can be completed in a single day. To further simplify the process, the starting combinatorial library will be consisting of DNA nucleic acid strands as opposed to RNA molecules. The rationale behind opting for DNA aptamers is primarily for the reason that DNA strands are far more stable than unmodified RNA strands [198]. Therefore, the synthesis of the DNA combinatorial library is simpler and more cost effective as modifications prior to selection to prevent nuclease degradation are not necessary as it would be if RNA aptamers were to be employed.

An obvious challenge upon selection and identification of the aptamers against the MUC1/Y splice variant is the apparent lack of *in vivo* stability of the unmodified aptamer and its fast renal clearance from the system. With the intention of improving the stability of the aptamers and improve its pharmacokinetic properties, the attachment of PEG would appear the most suitable approach, as the polymer possesses numerous advantages. Other than the functional benefits of PEG, the main advantage of using PEG would be that the polymer is not toxic and does not induce an immunogenic response, which coincides with the properties of the aptamers. With the aim to pegylate the selected aptamers, functional groups that need to be incorporated on the aptamer require careful consideration, so as to be reactive with the moieties available on PEG. The conditions which yield the greatest pegylated product need to be defined based on the reactive moieties present on the PEG and the aptamer. In addition to the conjugation process, it is essential to determine a method to purify the conjugated aptamer from the free aptamer and PEG. Also necessary is a bench top detection system that is rapid, simple to use and provides evidence for the presence of the conjugated aptamer relatively to the free PEG and aptamer without the need to use mass spectrometric analysis. Finally, subsequent to the modification and purification steps, vitally important is to establish any potential changes in the affinity of the aptamer for the target through several *in vitro* investigations before these aptamers can be considered further for *in vivo* studies.

CHAPTER TWO

MATERIALS AND METHODS

2.1 MUC1/Y PEPTIDES

The 10 amino acid (10mer) and 20 amino acid (20mer) MUC1/Y peptide sequences were designed by Dr Huma Khan in the aptamer group of the Open University. The 10mer peptide has a sequence of TEKNAFNSSL, with a molecular weight of 1110.17. The 20mer peptide has a sequence of SVPSSTEKNAFNSSLEDPST, with a molecular weight of 2097. The peptides were supplied HPLC purified by the Oligonucleotide and Peptide Synthesis Unit, Queen's Medical Centre, University of Nottingham (Nottingham, UK).

2.2 SINGLE STRANDED DNA LIBRARY & PRIMERS

Each DNA strand in the single stranded oligonucleotide combinatorial library used in the SELEX method was composed of a central 25 nt randomised region flanked by hybridising sequences comprised of 23 and 24 nt on the 5′ and 3′ end respectively. The hybridising sequence at the 5′ was GGGAGACAAGAATAAACGCTCAA and at the 3′ TTCGACAGGAGGCTCACAACAGGC. The amplification primers are 5′-GGGAGACAAGAATAAACGCTCAA (forward primer) and 3′ -GCCTGTTGTGAGCCTCCTGTCGAA (reverse primer). The primer sequences were selected based on previous research [197]. The combinatorial oligonucleotide library and the amplification primers were supplied HPLC purified by the Oligonucleotide and Peptide Synthesis Unit, Queen's Medical Centre, University of Nottingham (Nottingham, UK).

2.3 BIOTINYLATION OF MUC1/Y PEPTIDES

The MUC1/Y peptides (10mer and 20mer) were labelled with biotin-XX using a FluoReporter® Biotin-XX protein labelling kit. Each peptide (5.1 mg) was dissolved in 0.1 M sodium carbonate pH 8.4 (1 mL). A 20 mg/mL (35.2 mM) solution of biotin-XX SE was prepared by dissolving 8.0 mg of biotin-XX SE in DMSO (0.4 mL). The reaction was carried out according to the manufacture's protocol with the exception of using a 1:2 molar ratio instead of the recommended 1:10 molar ratio of peptide to biotin. 260.93 µl and 130.83 µl of the biotin-XX SE solution was added to the 10mer and 20mer peptide solution respectively. The mixtures were stirred at ambient temperatures for 2 hours. The reaction

mixtures were loaded on a PD-10 desalting column, using PBS (pH 7.2) as the eluant. The first 3 mLs of flow through were collected in 1 mL fractions and analysed by UV-Vis spectroscopy at an absorbance wavelength of 256 – 258 nm. The fractions were further purified by Reverse Phase HPLC (RP-HPLC) on a Waters 616 Millennium 32 system using a Waters Symmetry™ C_{18} column (4.6 x 250 mm). The absorbance was recorded on a Waters 996 Photodiode Array UV-Vis detector at the wavelength of 215 nm. The solvent gradient at a flow rate of 1 mL/ min was **0′:** A = 85% / B = 15%, **5′:** A = 75% / B = 25%, **25′:** A = 35% / B = 65%, **25.01′:** B = 100%, **30′:** B = 100%, **30.01′:** A = 85% / B = 15%, where A = 0.1% trifluoroacetic acid (TFA) in 2% acetonitrile (ACN), B = 0.1% TFA in 90% ACN.

Materials: FluoReporter® Biotin-XX protein labelling kit was purchased from Invitrogen Molecular Probes (Paisley, UK). PD-10 desalting columns were purchased from GE Healthcare (Chalfont St Giles, UK). Acetonitrile and trifluoroacetic acid (TFA) (HPLC grade) were purchased from Fisher Scientific UK (Loughborough, UK). Dimethyl sulfoxide (DMSO) (molecular biology grade) was purchased from Sigma-Aldrich (Dorset, UK).

2.4 PCR AMPLIFICATION OF THE SELEX LIBRARY

The PCR reaction was performed in a thin walled PCR tube as previously described [197]. The PCR amplification mixture was composed of 10^{15} ssDNA 72 mer SELEX library (10 μM), 5′ forward primer (250 μM), 3′ reverse primer (10 μM), 0.1 μl/μl of 10x PCR buffer containing $MgCl_2$ (40μl in 400μl final volume), 10 mM of dNTP mixture containing 25 mM of dATP, dCTP, dTTP, dGTP), $MgCl_2$ (0.675 mM), and Taq DNA polymerase (3 μl 5U/μl).

The following PCR cycle programme involves a total of 100 cycles. Each PCR cycle consisted of: • initial denaturation step (95°C, 10 min) • addition of first 1.5 μl Taq DNA polymerase • second denaturation step (95°C, 1.5 min) • annealing step (56°C, 0.5 min) • extension step (72°C, 1.5 min). • upon completion of the first 50 cycles, further 1.5 μl Taq DNA polymerase was added and a further 50 cycles were performed. • final extension step (72°C, 10 min). The final reaction solutions were stored at 4°C until used for selection.

Materials: For combinatorial library and amplification primer refer to section 2.2. 0.2 mL thin walled PCR tubes and magnesium chloride hexahydrate were purchased from Fisher Scientific UK (Loughborough, UK). PCR 10x buffer and Taq DNA polymerase were purchased from Sigma-Aldrich (Dorset, UK). dNTP's were obtained from Promega (Hampshire, UK).

2.5 DOUBLE STRANDED DNA PCR

To perform the double stranded PCR amplification of the ssDNA obtained from selection, a reaction mixture of the following composition was used: 50 µl of DNA from selection, 5' forward primer (1 µM), 3' reverse primer (1 µM), 10x PCR buffer containing $MgCl_2$ (0.1 µl/µl), dNTP mixture (0.5 mM), $MgCl_2$ (0.69 mM), and Taq DNA polymerase (1 µl 5U/µl).

The following PCR cycle programme involves a total of 50 cycles. Each PCR cycle consisted of: • initial denaturation step (95°C, 10 min) • addition of 1 µl Taq DNA polymerase • second denaturation step (95°C, 1.5 min) • annealing step (56°C, 0.5 min) • extension step (72°C, 1.5 min). • final extension step (72°C, 10 min). The final reaction solutions were stored at 4°C until further use. Refer to section 2.4 for materials.

2.6 IMMOBILISATION OF MUC1/Y PEPTIDES ON STREPTAVIDIN COATED TUBES

The biotinylated MUC1/Y peptides (10mer and 20mer) were immobilised onto the streptavidin coated tubes according to the manufacture's protocol. The biotinylated peptide solutions containing *ca* 30ng of either 10mer or 20mer peptide in PBS, pH 7.2 (300 µl) to give a final concentration of 65.54 nM and 41.29 nM respectively, were added to the streptavidin coated tubes. The solutions were incubated for 3 minutes at 37°C and removed after incubation. The tubes were washed with 300 µl binding solution and blot dried five times. The tubes were stored in binding solution at 4°C until used for selection.

Materials: Streptavidin coated PCR tubes were purchased from Roche Applied Science (Burgess Hill, UK).

2.7 *IN VITRO* SELECTION USING SALT GRADIENT ELUTION TECHNIQUE

The amplified single stranded DNA library (130 µl) was added to the tubes containing previously immobilised 10mer or 20mer peptides. After incubation at ambient temperature for an hour under gentle shaking, any excess library solution was removed and the tube was washed with binding solution (300 µl). A gradient elution of the bound aptamers was performed by washing the tube with NaCl solutions (50 µl) of increasing concentration starting from 0.2 M to 1.5 M elution solution using 0.1 M increments and incubating with gentle mixing for 1 minute. Each 50 µl washing obtained from the different concentration steps were desalted using microcons (3 kDa MWCO). The desalted 1.3 M, 1.4 M, and 1.5 M fractions were amplified as described in section 2.5 and analysed by 2 % agarose gel electrophoresis where the presence of aptamers was detected at 75 bp. Aptamers present in these fractions were selected for cloning.

1.5 M elution solution, pH 7.2: 1.5 M NaCl, 20 mM KCl, 5 mM $MgCl_2$. The pH was adjusted with 0.01 M NaOH. The incremental molarity of the elution solution is solely based on the concentration of NaCl. The concentrations of KCl and $MgCl_2$ remain constant throughout all the elution solutions. The elution solution was prepared as previously described in literature [197], with the addition of KCl.

Materials: Sodium chloride (laboratory grade), magnesium chloride hexahydrate, and microcons (3kDa, 10kDa, 30kDa MWCO) were purchased from Fisher Scientific UK (Loughborough, UK). Potassium chloride and sodium hydroxide pellets were purchased from Sigma-Aldrich (Dorset, UK).

2.8 *IN VITRO* SELECTION USING TEMPERATURE GRADIENT ELUTION TECHNIQUE

The incubation of the library to the immobilised peptides was performed as described in section 2.7. After incubation, any excess library solution was removed and the tube was washed with binding solution (300 µl). The aptamers were eluted with successive washings using sterile water (50 µl) at increasing temperatures from 25°C to 95°C, increasing by 5°C increments with gentle mixing for 2 minutes for each step. The washings obtained from the 85°C, 90°C and 95°C were amplified as described

in section 2.5 and were analysed by 2 % agarose gel electrophoresis where the presence of aptamers was detected at 75 bp. Aptamers present in these fractions were selected for cloning.

2.9 AGAROSE GEL ELECTROPHORESIS

The 2% (w/v) agarose gel was prepared by adding 2g of agarose to 1x TBE (100 mL) solution. The solution also contained the DNA stain ethidium bromide (5 µl of 10 mg/ mL solution). The samples (10 µl) containing 6x loading dye (2 µl) were loaded into gel wells for analysis. A 25bp or 100bp DNA step ladder (5 µl) containing 6x loading dye (2 µl) was loaded into a separate gel well for comparing band size present in the samples. The electrophoresis was performed at 100 V in 1x TBE running buffer. The gel was imaged on the Kodak digital image station 440CF.

1x TBE (Tris-Borate EDTA): Trizma base (5.4 g), boric acid (2.74 g) and EDTA (0.46 g) were dissolved in UHP water (500 mL).

Materials: Agarose was purchased from Fisher Scientific UK (Loughborough, UK). Boric acid, trizma® base, ethylenediaminetetraacetic acid (EDTA) and ethidium bromide were purchased from Sigma-Aldrich (Dorset, UK). 25bp DNA step ladder, 100bp DNA step ladder, and 6x loading dye were purchased from Promega (Hampshire, UK).

2.10 CLONING AND SEQUENCING

The pool of aptamers amplified from sections 2.7 and 2.8 were cloned into TOPO vectors, using the TOPO TA cloning kit, according to the manufacture's protocol. Each transformation mix (10 µl and 50 µl) was spread onto the LB agar plates and left to incubate at 37°C overnight. The colonies from each fraction were PCR amplified as described in section 2.11 and analysed by 2 % agarose gel electrophoresis. The presence of aptamers was detected at 275 bp. The positive clones were sent for sequencing to Macrogen, Korea.

Luria Bertani (LB) Agar plates: The LB medium contained tryptone (5g), yeast extract (2.5 g), NaCl (2.5 g), Agar (7.5 g) and UHP water (500 mL). The solution was autoclaved and once cooled to *ca* 50°C,

500 µl of ampicillin (50 mg/mL) was added to the medium which was immediately poured into the plates. The prepared LB agar plates were stored at 4°C until further use.

Materials: TOPO TA cloning® kit was acquired from Invitrogen molecular probes (Paisley, UK). Tryptone, yeast extract, agar, and ampicillin were purchased from Sigma-Aldrich (Dorset, UK). Sodium chloride (laboratory grade) was purchased from Fisher Scientific UK (Loughborough, UK). Petri dishes (100mm x 15mm) were purchased from PAA Laboratories (Somerset, UK).

2.11 PCR OF POSITIVE CLONES

Each clone was added to a PCR mixture with the following composition of M13 forward primer (0.25 µM), M13 reverse primer (0.25 µM), 10x PCR (0.1 µl/µl) buffer containing $MgCl_2$, dNTP mixture (0.5 mM), $MgCl_2$ (0.69 mM), sterile water (15.98 µl) and Taq DNA polymerase (1 µl 5U/µl).

The following PCR cycle programme involves a total of 50 cycles. Each PCR cycle consisted of: • initial denaturation step (95°C, 10 min) • addition of 0.2 µl Taq DNA polymerase • second denaturation step (95°C, 1.5 min) • annealing step (56°C, 0.5 min) • extension step (72°C, 1.5 min). • final extension step (72°C, 10 min). The final reaction solutions were stored at 4°C until further use.

Materials: Forward and reverse M13 primers purchased from Invitrogen life technologies (London, UK). For other materials refer to section 2.4.

2.12 PEGYLATION OF APTAMERS

Linear 32 kDa m-PEG (8.464 mg, 259.49 nmol) dissolved in 170 µl of dimethylformamide (DMF) was added in 1 mg portions every 20 minutes to a solution of 3′ C_6 amino modified aptamer (12 µl, 6 nmol) in 30 µl of sodium carbonate pH 10. The reaction was performed at 60°C under gentle shaking. After the addition of the final 1 mg portion of PEG, the mixture was mixed for a further 1 h at 60°C. The mixture was lyophilised using speed vacuum GenevacSF50. The solid residue was reconstituted in 170 µl of sterile water for purification by anionic exchange HPLC on a Waters 616 Millennium 32 system, using a

Mono Q® HR 5/5 1 mL column (10 μm). The absorbance was recorded on a Waters 996 Photodiode Array UV-Vis detector at the wavelength of 260 nm. The elution gradient at a flow rate of 1 mL/ min was **0′:** A = 40% / B = 60%, **20′:** A = 80% / B = 20%, **21′:** A = 100%, **25′:** A = 100%, **26′:** A = 40% / B = 60%, **39′:** A = 40% / B = 60%, where A = 1.8 M ammonium formate pH 8, B = Water. The purified conjugated aptamer was lyophilised and analysed by 2% agarose gel electrophoresis.

53.4 mM Sodium carbonate, pH 10: Na_2CO_3 (0.566 g), $NaHCO_3$ (0.391 g) in a final volume of 100 mL UHP water and filtered using stericup filter units.

1.8 M Ammonium formate, pH 8: Formic acid (135.86 mL) in a final volume of UHP water (2 L). pH adjusted with concentrated ammonia and filtered using Empore C_{18} disk.

Materials: NHS-modified 32kDa polyethylene glycol was purchased from NOF Corporation (Tokyo, Japan). 3' C_6 amino modified 25 nt and 3' C_6 amino modified- 5' Cy3 labelled 25 nt oligonucleotides were purchased from Integrated DNA Technologies (Budapest, Hungary). Mono Q HR anionic exchange column was purchased from GE Healthcare (Chalfont St Giles, UK). *N,N,* Dimethylformamide (DMF) (molecular biology grade) and empore C_{18} 47 mm disks were purchased from Sigma-Aldrich (Dorset, UK). Sodium carbonate, sodium hydrogen carbonate, formic acid (98%), and 35% ammonia solution were purchased from Fisher Scientific UK (Loughborough, UK). Stericup Filter Units were acquired from Millipore (Durham, UK).

2.13 NATIVE POLYACRYLAMIDE GEL ELECTROPHORESIS

To prepare polyacrylamide gels using 29:1 acrylamide : bisacrylamide, the volume of acrylamide required in the mixture was determined by the formula: % needed / % acryl:biacryl (30%) * final volume (mL). An 18 % native PAGE was used to analyse aptamer samples. A 10 mL solution contained 30% acrylamide (6 mL), 5x TBE (2 mL), UHP water (1.87 mL), 10% (w/v) ammonium persulfate (110 μl) (APS) and TEMED (20 μl). The APS and TEMED solutions were added to the mixture immediately prior to pouring the gel solution for gel formation. The samples (10 μl) containing 6x loading dye (2 μl) were loaded into gel wells for analysis. A 25bp or 100bp DNA step ladder (5 μl) containing 6x loading

dye (2 µl) was loaded into a separate gel well for comparing band size present in the samples. The electrophoresis was performed at 100 V in 1x TBE running buffer. The gels were post stained in 100 mL of 1x TBE containing the DNA stain ethidium bromide (10 µl of10 mg/ mL solution) for 10 minutes. The gel was imaged on the Kodak digital image station 440CF.

TBE (Tris-Borate EDTA) (5x): Trizma base (13.5g), boric acid (6.86g) and EDTA (1.16g) were dissolved in UHP water (250 mL).

Materials: acrylamide / bisacrylamide 30% solution, *N,N,N',N'*-Tetramethylethylenediamine (TEMED), and ammonium persulfate (APS) were purchased from Sigma-Aldrich (Dorset, UK). For other materials refer to section 2.9.

2.14 BARIUM CHLORIDE-IODINE STAIN FOR GELS TO VISUALISE PEG

In order to visualise either free or conjugated PEG, either the agarose or the native polyacrylamide gels were post stained in 0.1 mol/L barium chloride solution (100 mL) for 20 minutes at ambient temperature with gentle shaking, followed by the gradual addition of 0.05 mol L^{-1} iodine aqueous solution (10 mL). The gels were further soaked in the barium chloride-iodine solution with gentle shaking until brown stained bands of PEG were visibly observed.

Materials: 0.1 mol/L barium chloride solution and 0.05 mol L^{-1} iodine aqueous solution were purchased from Sigma-Aldrich (Dorset, UK).

2.15 APTAMER AFFINITY CHROMATOGRAPHY

The MUC1/Y peptides, 10mer (10.76 µmol) and 20mer (10.04 µmol) were individually immobilised to the NHS-activated 1 mL HiTrap sepharose columns according to the manufacturer's instructions. Depending on the aptamer, the concentrations of the solutions to analyse were in the range of 1.17 µM – 4.27 µM having an optical density of 1.0. Each 1 mL aptamer in binding solution was added to the column and incubated for 1 h at ambient temperature. The column was then eluted with the elution solutions (1 mL) of increasing NaCl concentrations ranging from 0.2 M – 1.6 M, at 0.2 M increments.

The column was finally washed with a 3 M sodium thiocyanate solution (1 mL x 3), followed by binding solution (1 mL x 3), to regenerate the column. Each eluted fraction was analysed by UV-Vis spectroscopy measuring the absorbance at 260 nm on a UVkon XL spectrophotometer (Bio-Tek instruments).

Binding solution, pH 7.2: 100 mM NaCl, 20 mM KCl, 5 mM $MgCl_2$. The pH adjusted with 0.01 M NaOH.

Materials: Unmodified 72 nt oligonucleotides were purchased from MWG (Ebersberg, Germany). Unmodified 25 nt oligonucleotides were purchased from Integrated DNA Technologies (Budapest, Hungary). HiTrap NHS-activated affinity column was purchased from GE Healthcare (Chalfont St Giles, UK). Sodium thiocyanate and potassium chloride was purchased from Sigma-Aldrich (Dorset, UK). Sodium chloride (laboratory grade) and magnesium chloride hexahydrate were purchased from Fisher Scientific UK (Loughborough, UK).

2.16 ELECTROPHORETIC MOBILITY SHIFT ASSAY (EMSA)

Each aptamer was incubated with the 10mer and the 20mer peptides in a 1:1000 ratio. Aptamer (2 μM) and either 10mer or 20mer peptide (2000 μM) was incubated in binding solution. The incubation was performed with two different conditions: **1)** For 1 h at ambient temperature under gentle shaking **2)** at 95°C for the first 5 minutes and then for 1 h at ambient temperature under gentle shaking. After incubation, the samples were analysed by native PAGE at ambient temperature at 100V in 1x TBE running buffer. 14 – 16 %, 18 % and 7 % native polyacrylamide gels were used for resolving 75nt, 25nt and pegylated 25nt long aptamers respectively. The bands were visualised as described in sections 2.13 and 2.14. For materials refer to sections 2.13 and 2.15.

2.17 APTAMER THERMAL DENATURATION & RENATURATION

A mixture composed of the aptamer (0.180 μM), and either the 10mer or 20mer peptide (180 μM) in a 1: 1000 ratio aptamer to peptide in binding solution was incubated for 1 h at ambient temperature.

Subsequently, the 10x SYBR green I nucleic acid gel stain (2.5 µl) was added to the reaction mixture which was kept on ice (4°C) for 30 minutes. The aptamer-peptide mixture was heated from 3°C to 100°C at a 1°C/ min rate (denaturation step), followed by a cooling down process from 100°C to 3°C at a 1°C/ min rate (renaturation step). The fluorescence emission was recorded every minute during the denaturation step and every 10°C increments from 80°C to 10°C and thereafter, at 5°C and 3°C. The denaturation-renaturation cycles were made in triplicates and ran on a DNA Engine Opticon 2 Real Time PCR (MJ Research) at 520 nm using excitation at 497nm. Final analysis of the data was carried out using the software Origin 6.0.

Materials: SYBR green I nucleic acid gel stain was purchased from Sigma-Aldrich (Dorset, UK).

2.18 CELL CULTURE

MCF-7, DU145, CALU-6 and A498 cells were grown at 37°C and 5% CO_2 in minimal essential medium eagle (MEM) cell culture medium supplemented with 10% foetal bovine serum and 1% L-glutamine.

Materials: DU145 (human prostate carcinoma), MCF-7 (human mammary gland carcinoma), Calu-6 (human lung carcinoma) and A498 (human kidney carcinoma) cell lines were provided by Antisoma Research Limited (Hertfordshire, UK). Trypsin EDTA, minimal essential medium eagle (MEM), L-Glutamine, foetal bovine serum and cell culture flasks (75 cm^2, surface treated with filter cap) were purchased from PAA Laboratories (Somerset, UK). Hanks balanced salt solution was obtained from Sigma-Aldrich (Dorset, UK).

2.19 FLUORESCENCE ACTIVATED CELL SORTING (FACs)

Cy3 fluorescently labelled aptamers (2 µM) were incubated with MCF-7, DU145, CALU-6 and A498 (400,000 cells) in 200 µl of PBS at 37°C for 1 h. Upon incubation, the cells were washed with PBS (400 µl x3) and resuspended in PBS (400 µl). The analysis was performed on FACSCalibur, Becton Dickinson flow cytometer with a maximum sort rate of 10,000 total cells / second.

Materials: 5' Cy3 labelled 25 nt oligonucleotides were purchased from Integrated DNA Technologies (Budapest, Hungary).PBS tablets were purchased from Sigma-Aldrich (Dorset, UK). 5 mL FACs polystyrene tubes (12 mm x 75 mm) were purchased from Greiner Bio-One LTD (Stonehouse, UK). For other materials refer to section 2.18.

2.20 FLUORESCENCE MICROSCOPY

MCF-7, DU145, CALU-6 and A498 cells (50,000 cells) were plated in duplicates on round glass inserts, placed inside each well of a 24 well cell culture plate and grown overnight. The adhered cells were washed with PBS (100 µl), followed by incubation with aptamer (0.5 µM, 150 pmol) in serum free MEM (300 µl) per well for 15 min, 1h and 3 h at 37°C. Subsequently, the cells were washed with PBS (200 µl x3), followed by the addition of 4% paraformaldehyde (PFA) (200 µl) and incubated for 15 min at ambient temperature. The cells were then washed with PBS (200 µl x2), treated with 0.1 % triton (200 µl) for 10 min at ambient temperature, followed by a PBS wash (200 µl x2). The cells were then incubated with Hoechst 33258 (75 µl of 1 mg / mL) for 15 min at ambient temperature, followed by a PBS wash (200 µl x3). The glass inserts were mounted on to glass microscope slides using CFPVOH poly vinyl alcohol and AF100 anti-fadent (10:1 ratio). The slides were visualised under a BX61 motorised Olympus DP72 fluorescence microscope digital camera, using Cy3 and DAPI filter with 40x magnification. The images were analysed using the ImageJ software.

Materials: 24 well tissue culture plate, flat bottom were purchased from PAA Laboratories (Somerset, UK). Round glass coverslips (12 mm) and glass microscope slides (76 mm x 26 mm) were purchased from Fisher Scientific UK (Loughborough, UK). Paraformaldehyde (PFA), triton x100, and bisBenzimide H33258 were purchased from Promega (Hampshire, UK). CFPVOH (Poly vinyl alcohol) and AF100 (Antifadent) were purchased from Citifluor LTD (London,UK).

CHAPTER THREE

IN VITRO SELECTION

3.0 INTRODUCTION

Since the development of the systematic evolution of ligands by exponential enrichment (SELEX) process, numerous aptamers have been generated against various different selected targets. Typically the conventional method of SELEX employs the use of affinity chromatography columns and magnetic beads for the immobilisation of the target or nitrocellulose membrane for ultrafiltration. Successive purification steps are performed of the nucleic acid based combinatorial library against the selected target. However, as the demand for aptamers is rising in the area of molecular recognition, diagnostics and therapeutics, prerequisites of improving the standard selection technology has become essential. Many researchers have developed new alternative methods which are based on the core principles of SELEX for the discovery of aptamers possessing high affinity and specificity for their target. Moreover, the new established selection techniques aim to offer a much faster, less laborious, low reagent consuming, and hence cost effective systems.

3.1 NON-CONVENTIONAL SELECTION METHODS

New variations to the classical SELEX system have been approached by various researchers to overcome many of the existing issues such as the laborious and time consuming process, along with utilisation of large reagents quantities, thus increasing the costs of the entire process.

A new one-step selection method, MonoLex has been applied to generate high affinity binding DNA aptamers to a Vaccinia virus, which was employed as model organism representative of complex targets. The MonoLex system was developed to eliminate repetition of the successive rounds of selection and minimize the process to a single affinity chromatography with a final amplification step. The selection entails a combinatorial library consisting of oligonucleotides 64 bases long of which, the central randomised region is comprised of 20 nucleotides. Primarily the combinatorial library was subjected to the affinity capillary column that was coupled to virus free cell culture supernant, preventing any non-specific binding. Subsequent to the negative selection, the unbound library was applied to the affinity column per-coupled to the heat inactivated virus particles. The low binding oligonucleotides were

removed by repeated buffer washes, preparing the column to be sliced into many segments. The target bound oligonucleotides captured within each segment were released by denaturation of the target, followed by quantitative real-time PCR amplification. Results showed that in each segment more than one species of aptamers were present, hence, the amplified mixture was referred to as "polyclonal" aptamers. Although the cloned and sequenced aptamer pool of interest showed no sequence homology, *in vitro* binding studies demonstrated the aptamers obtained from the selection having great affinity and specificity for the target virus. The binding of the target to the aptamer could be monitored and detected at extremely low concentrations of target and aptamer, 10^9 particles/ mL and 10pM respectively. The studies also provided evidence of aptamer specificity, with the aptamer binding only to the virus particles and not to the control particles of a similar size. This result was also consistent, as non-specific binding was not detected even when repeated with higher aptamer concentrations. In further investigations involving three different cell types, the biological activity was inhibited and spread of infection was prevented at an aptamer concentration of 2.5μM.

The MonoLex system has demonstrated that aptamers possessing high affinity and specificity can be successfully generated against a virus, which is a realistic representative of a complex target model. An additional benefit this system brings is that competition between oligonucleotides of similar or differing affinities during the PCR amplification step is dramatically reduced, thus, significantly improving the identification of high binding aptamers for the chosen target [199].

The cell-based systematic evolution of ligands by exponential enrichment (cell-SELEX) is a method used to generate aptamers capable of identifying cell surface biomarkers present on whole living cells. The key advantage cell-SELEX compared to many other selection methods is that details of the biomarker does not need to be known prior to selection and multiple aptamers can be generated against different cell surface biomarkers simultaneously. The oligonucleotide library employed in this selection incorporated a central random region of 45 nucleotides flanked by 20 bases long primer region. The PCR amplification primers were modified, where the 5' primer was labelled with a fluorescent tag, fluorescein isothiocyanate (FITC), and the 3' primer was triple biotinylated. A double stranded (ds) PCR

amplification was performed on the combinatorial library and incubated with streptavidin coated beads. Succeeding this, the dsDNA molecules were denatured and the FITC labelled single stranded oligonucleotides were separated from the biotinylated ssDNA strands. The fluorescent labelled ssDNA pool was incubated with whole cells, followed by the removal of majority of the library representing the non-binding selection. The bound aptamers were eluted by thermal denaturation and amplified by double stranded PCR. The dsDNA was separated to ssDNA as mentioned above for the next round of selection. However, between each selection round, flow activated cell sorting (FACs) was performed to ensure the enriched library of oligonucleotides is gradually showing greater binding and specificity compared to the starting combinatorial library. Throughout the successive rounds of selection, stringency of the process was progressively improved by extending the wash time, increasing the wash volume as well as the total number of washes. Aptamers at the final 23rd round of selection were sequenced and those which revealed sequence homology were taken further for investigation. The results demonstrated that with each enriching round of selection, the binding affinity of the evolving aptamers increased. This was suggested by the amplified fluorescence intensity observed during flow cytometric analysis, which differed from that observed in the control cell line (CCRF-CEM). The selected aptamers possessed binding affinities ranging from nanomolar to the picomolar Kd values. However, also present in the final panel of aptamers, were a minority of sequences which displayed greater interaction with the CCRF-CEM cells compared to the positive cell line (Ramos). This proposed the likelihood of greater expression of a particular target present on the CCRF-CEM cells than on the Ramos cells. Thus, increasing the stringency of the wash steps during the selection rounds assisted in achieving aptamer interaction with biomarkers expressed at low levels. The selected aptamers also displayed target specific binding in real human biological samples (bone marrow mixture) and were able to distinguish between normal and diseased cells. This selection system did not entail a counter selection, as the aim was to generate aptamers against as many different receptors as possible on the surface of target cells and is an efficient starting point to discovering biomarkers existing on the surface of diseased cells [200].

Prior to the cell-SELEX described above, Weihong Tan's group had also performed a cell-SELEX method, which had incorporated a counter selection within the protocol. However, in this selection, the

target cell line was exchanged with the control cell line and *vice versa*. Thus, the targets cells were the CCRF-CEM cells and the control cells were the Ramos cells. The protocol followed steps as described above with the exception that upon incubation of the library with the CCRF-CEM cells, the bound sequences were eluted and subjected to 5-fold excess of Ramos cells. The unbound aptamers to the negative cell line were amplified for another round of selection. A total of 20 rounds of selection were performed in this method and the enriching library was monitored by flow cytometry at the end of each round of selection. The counter selection aided in reducing the number of aptamer sequences that bound to the Ramos cells. Thus, the Ramos displayed no sizeable shift in the fluorescence intensity, though the fluorescence intensity increased with each progressive round of selection with the CCRF-CEM cells. This resulted in a pool of DNA molecules possessing high specificity for receptors expressed on the cells surface [142]. A similar selection was also demonstrated by Blank et al [143], performing SELEX employing a counter selection and screening each successive pool by FACs analysis. The incorporation of flow cytometric analysis in the SELEX process has allowed each pool of DNA molecules obtained at the end of each round of selection to be analysed, providing a clear indication of the potential interaction which may be taking place between the aptamer and the target. Furthermore, the benefit of employing flow cytometry is that as the binding affinity and specificity of the aptamers for their targets increases, this is effectively demonstrated by the increase in the fluorescence intensity. Hence, the changes in the fluorescence intensity are an accurate display of evaluating at each round of selection if further rounds of selection are necessarily required.

In most of the well known traditional SELEX protocols, a combinatorial library typically consists of primer hybridising sequences flanking either side of the unknown random domain. By and large the primer sequences comprise fifty percent if not more of the library sequences. Hence, it has often been suggested that primer sequences have the potential to interfere in the binding of the unknown region with the target. Typically, the final pool of aptamer sequences can consist of a section of the primer sequence which is either involved in base pairing with the random region to adopt secondary structures or is involved in the selected binding sequences. This form of involvement of the primer sequences can lead to non-specific binding and can produce false positive binding sequences with relevance to the unknown

random region. Therefore, to eliminate the compromise on the binding potential of the randomised region to the target alone, two novel selection methods have been devised. One method entails a combinatorial library consisting of minimal primers (MP) flanking the central unknown domain and the second method involves a primer free (PF) library of oligonucleotides. The initial library for both protocols is double stranded by PCR with primer sequences flanking the central variable region of 27 nt long for the MP library and 30 nt for the PF library. The MP library contains an endonuclease site at the 5′ end for an enzymatic cleavage to take place, where the enzyme is capable of recognising the dsDNA but will only cleave one strand. An endonuclease restriction site is also present at the 3′ end to ensure cleavage of the dsDNA substrate. The cleavage at the 5′ and 3′ regions leaves adjoined 2 nt (dock sites) of the primer sequences at either sides of the unknown random region, resulting in a ssDNA library of 31 nt long. The cleaved library was gel purified and subjected to melanoma cells for selection. The dock sites of the bound sequences were hybridised with their respective 5′ and 3′ end bridge pair sequences and ligated to regenerate the primer regions, which were previously removed through enzymatic cleavage. The ligated products were amplified by PCR and this entire process was repeated for each additional round of selection. The second protocol involving the PF library contains an endonuclease site at both the 5' and the 3′ end, where the endonuclease can recognise the dsDNA but will only cleave one strand. This primer free approach explores two different sequence based libraries: in one library, the cleavage takes place at the 3′ end leaving adjoined 2 nt (dock sites) of the primer sequence to the unknown random region, resulting in a ssDNA library of 32 nt long. The second library only contains the 30 nt unknown random regions only, as both primers from the 5′ and 3′ end had been enzymatically cleaved. Both libraries were subjected to selection and as with the MP library method, the bound aptamers were obtained for regenerating primer regions through a hybridisation and ligation processes. The ligated products were amplified by PCR and these processes were repeated for each additional round of selection [201].

This SELEX system, which employs either the minimal primer or primer free library, also includes additional steps such as enzymatic cleavage of primer, primer regeneration and purifying the amplified library via polyacrylamide gel electrophoresis, which are lacking in a typical SELEX method.

Nonetheless, this method provides a different approach for incorporating primers for sequence amplification without their interference in the binding of the unknown random sequences to the target. An additional benefit of this system is that the necessissity to reduce the total length of the aptamer sequence is eliminated, whereas truncation would be essential if a typical library containing full primer sequences was employed.

Aptamer selection has also been performed using capillary electrophoresis (CE) integrating the classical principles of SELEX. In the CE-SELEX method, the binding sequences are separated from the non-binders via electrophoresis. This method allows incubation of the target with the combinatorial nucleic acid library to take place in free solution, removing the prerequisites of stationary supports. The incubated mixture is then separated under high voltage through a capillary, whereby the non-binders travel through the capillary at the same mobility, in spite of their sequence or length. However, the binding sequences bring about a mobility shift as, upon binding to the target, a change in the size and charge occurs. Based on this mobility shift, the binding sequences are collected for PCR amplification to be taken into further rounds of selection. This technique has been demonstrated by successfully obtaining aptamers against the human IgE with nanomolar affinity. To obtain the IgE aptamers, a library which contained sequences composed of two primer regions of 20 nt flanking a 40 nt unknown random region was incubated with the target at room temperature for 30 minutes. The incubated samples were injected onto the capillary with a current applied of ~ 70 μA during separation, maintaining the capillary column temperature at 25°C during electrophoresis. The unbound and bound sequences flowing through the column were monitored by UV detection at 254 nm. The non-binding DNA sequences, which are negatively charged, migrate swiftly through the capillary in CE buffer, pH 8.0, as a single band. However, the migration of the bound DNA sequences is decelerated due to the complex formation with the IgE, which is a large molecule and is slightly positive in the CE buffer. The CE fractions of the bound sequences were PCR amplified for further rounds of selection. A total of four rounds of selection were performed before the sequences were cloned. The nanomolar affinity of the sequences obtained compared to the aptamer sequences obtained for IgE using the conventional SELEX method consisting typically of 8-15 rounds of selection. CE-SELEX can be completed in as short as 2 days. However, this

method has a limitation, whereby the system is not applicable for all types of target molecules. For example, aptamers can only be generated in the CE-SELEX method against target molecules with a minimum size of ~ 4 kDa, as the target needs to be relatively large to cause a mobility shift upon binding to the DNA molecules [202,203].

Another selection process adapted from the CE-SELEX method is known as the non-equilibrium capillary electrophoresis of equilibrium mixtures (NECEEM). NECEEM is referred to as a non-SELEX selection method for generating aptamers against a known target, as this process only involves the partitioning steps and eliminates the necessity to amplify between each round of selection. It has been observed that only three partitioning steps were necessary with the NECEEM system to enhance the affinity of the DNA library for the target by beyond 4 orders of magnitude. The entire selection process takes 1 hour to complete, which typically entailed incubating the library with the target in free solution and upon incubation the mixture is injected into the capillary. Upon injection, voltage is applied to separate the bound sequences from the unbound as in the CE-SELEX system. However, the order of separation differs with the NECEEM, as the system is set up to allow separation to take place in the following order, eluting firstly with: (1) unbound free target, (2) bound target dissociated from the complex, (3) undissociated DNA-target complex, (4) DNA dissociated from the complex and finally (5) unbound free DNA. The NECEEM system was demonstrated by employing h-RAS protein as a target with a molecular weight of 21 kDa. The DNA library which was incubated with the h-Ras protein consisted of a random region of 39 nt flanked by two primer regions of 19 and 22 nt long and with the 5′ end of the DNA sequences fluorescently labelled. The incubated mixture containing the protein and the library was injected into the capillary and electrophoresis was performed applying an electric field of 600 V/cm in a buffer at pH 8.0. The eluting fractions containing the DNA sequences were detected by fluorescence at 520 nm and the protein was detected by UV at 280 nm. The fraction containing the target-DNA complex was further incubated with additional protein and subjected to another CE separation as in the first round. This process was repeated for a total of three rounds of partitioning, followed by separation of the aptamers from the complexes for a final PCR amplification and cloning.

The aptamers obtained via the NECEEM method had affinities of 0.3 μM which were comparable to aptamers that were obtained from the classical SELEX method of 0.6 μM [204,205].

Amongst many of the modified selection systems which are capable of reducing the total period of time taken to complete the selection process, NECEEM offers a further advantage, whereby PCR amplification is not required between each round of selection. Thus, this can be beneficial for employing non-amplifiable libraries consisting of modified bases and, furthermore, greater part of this system can be automated, thus producing an effortless aptamer selection system.

Another one step selection method was developed to save time and avoid the labour intensive process involved in the traditional SELEX system is referred to as the Aptamer Selection Express (ASExp) [206]. This technique employs three different combinatorial ssDNA libraries consisting of 10^{15} oligonucleotide sequences. The library consisted of a central random region of 30 nt, 40 nt or 50 nt, which was flanked by hybridising primer sequences composed of 20 nt for PCR amplification. The ssDNA library was amplified to generate a dsDNA library, which was ethanol precipitated and subsequently resuspended in PBS containing Mg^{+2} and Ca^{+2}. To this resuspended solution, the target (botulinum neurotoxin type A or B, light chain) was allowed to incubate with the library for 30 minutes, which causes the denaturation of the dsDNA, resulting in a formation of an ssDNA-target complex. Following incubation, magnetic beads covalently attached to random ssDNA (60mer) was added to the solution containing the ssDNA-target complex, which further forms a complex of ssDNA-target-magnetic bead-ssDNA. This complex is separated from the unbound dsDNA library using a magnet, whereby the supernatant containing the unbound dsDNA library is removed and the complex is thoroughly washed with water. Finally the bound aptamers were PCR amplified, cloned and sequenced. Although the ASExp system was in the early stages of development, the entire selection process was completed in 10 hours thus, accomplishing the original goal of developing a system which could generate aptamers by eliminating many of the standard time consuming process of the conventional SELEX system.

Aside from aiming to reduce the number of processes involved in the traditional SELEX method, some researchers have also aimed at adapting the separation technique involving the partition of aptamers bound to the target from the unbound oligonucleotides. Other than the typical separation processes, such as magnetic beads or some form of chemical treatment, electrophoretic separation is also seen as alternative for nucleic acids, proteins and their complexes. Gel electrophoresis allows separation to take place based on size and charge of each component, which will migrate across the gel at varying speeds. This approach was employed in an *in vitro* selection of RNA-based inhibitors of the human neutrophil elastase (hNE). In this selection process, an RNA library was utilised of which each strand in the library is modified with a splint oligonucleotide. A splint oligonucleotide is a DNA strand which is complementary to the hybridising primer sequence at the 5′ end of the RNA strand in the library and this splint oligonucleotide was conjugated at the 5′ end to a reactive valyl phosphate moiety (valP) via an amino linker, referred to as a DNA: valP complex. The DNA: valP complex was annealed to the RNA strands in the library (RNA: DNA: valP complex), which then followed by incubating the library complex with the target (hNE) for 5 – 10 minutes at 37°C. The resulting complex (RNA: DNA: valP: hNE) was resolved from the unbound complex (RNA: DNA: valP) on a 4% polyacrylamide gel. From the gel, the target containing complex (RNA: DNA: valP: hNE) was removed and the RNA was extracted through a crush and soak method. The recovered RNA was PCR amplified and subjected to another round of selection. In total, ten rounds of selection were performed before cloning and sequencing [207]. This principle of separation was also employed in other similar processes [208-210] for the partitioning of oligonucleotide-target complexes from the unbound oligonucleotides.

Amongst other typical methods of separation such as affinity chromatography and membrane filtration and those not so typical mentioned above, another method for separation is the surface plasmon resonance (SPR) technique. This biosensor technology is typically used to monitor molecular interactions between two or more molecules in real time and has been used to monitor interactions between different biomolecules such as nucleic acids, proteins, and peptides [211,212]. This technique has also been employed in aptamer selection against the hemagglutinin (HA) protein of the human influenza virus. The selection process entailed immobilisation of the HA protein to the NHS activated sensor chip, followed by

injection of the RNA library into the flow cell for 2 minutes. The RNA library contained 10^{13} different sequences with a randomised region of 30 nt. Subsequently to the injection of the library, binding buffer was injected and different fractions were collected. The bound aptamers were removed by ethanol precipitation and PCR amplified to enter another round of selection. In total five rounds of selection were performed. Aptamers obtained from this system possessed sequence consensus and displayed high affinity for the target HA protein [213]. In addition to employing the SPR technique in separation, this system has been shown to be valuable in the enrichment of binding aptamers from an oligonucleotide pool [214]. The advantages of employing this technique is that once the target is immobilised onto the sensor chip, the target can be repeatedly used, thus requiring the minimum quantity of the target. Furthermore, along with achieving separation of the molecules, affinity of the aptamers can simultaneously be obtained during the selection process.

Of all possible selection methods available, the most revolutionised system is of the microarray analysis, whereby high throughput screening of aptamers can be performed. On a single small glass chip, thousands of oligonucleotides of different sequences can be allocated in specific positions to which the chosen target can be incubated with and through fluorescence imaging the specific sequence can be identified from the emitting signal. An example of aptamer selection employing this techniques was demonstrated by Collet et al [215], where aptamers were screened against the lysozyme protein. Each 80 nt long biotinylated RNA from the aptamer template obtained from a previous selection (12^{th} round), was spotted on streptavidin coated slides and incubated for 30 minutes. Upon incubation, blocking solution was incubated for an hour, which was then followed with 30 minutes of incubation with the fluorescently labelled target protein. Subsequent to target incubation, the slides were washed with blocking solution to remove the unbound protein. After the washing step, the slides were stained with an oligonucleotide stain and then analysed. For the labelled protein, the limit of detection was 1 pg/ mL. The aptamer microarray analysis has also been thoroughly described by Ellington and coworkers [216,217]. The advantage of this technique dramatically reduces the overall time required to generate aptamers against a target, by eliminating the cloning and sequencing steps. Furthermore, this entire system can be automated, thus it is not labour intensive and requires minimal volume of target and oligonucleotides.

3.2 THE SIMPLEX METHOD

Upon taking into consideration the basic principles of the conventional SELEX along with the many advantages of the modified selection systems, a one-round selection system, termed "SimpLex", based on the original SELEX methodology had been designed within the group [197] with the aim to achieve high affinity binding aptamers. SimpLex is a simple, unique, one-round selection intended to serve the purpose of providing aptamers against the MUC1/Y peptides using minimal quantity of the combinatorial oligonucleotide library and target, in combination with providing efficiency pertaining to time and cost. The distinctive feature of this system is the elimination of the several rounds of amplification and enrichment of the library. With these stages avoided, each progressive elution step is potentially conditioned to facilitate the process towards acquiring high affinity aptamer(s) for the target ligand.

The SimpLex method, like most SELEX systems, utilises a synthetic single stranded (ss) DNA pool comprised of a 10^{15} random oligonucleotide sequences. Each oligonucleotide encompasses a unique central randomised nucleotide region which provides great variety of secondary DNA structures and is flanked either side by fixed primer sequences essential for PCR amplification. Prior to selection, the target (MUC1/Y peptides) was required to be modified via conjugation to biotin, for immobilisation to streptavidin coated PCR tubes. The amplified aptamer library is incubated with the pre-immobilised biotinylated peptides, allowing innumerable distinctive three-dimensional shaped aptamers to selectively interact with differing affinities for the biomarker. It is expected that the weak binding aptamers will be separated from the strong binders via a step elution gradient with either a sodium chloride based salt solution or a temperature gradient system. The aptamers obtained from the highest salt concentration and/or temperatures in the final selection step are PCR amplified for cloning and sequencing. Sequencing identifies the exact nucleotide arrangement of the randomised region for each aptamer, from which potential secondary structures can be determined using an oligonucleotide structure predicting programme. Prediction of secondary structures can aid in ascertaining the possible trends in the formation of particular structures that may be involved in binding.

3.3 MUC1/Y PEPTIDES: 10MER & 20MER

As the MUC1/Y splice variant is a transmembrane protein, the full protein cannot be utilised in the selection procedure, due to the complexities involved in its isolation and purification. Thus, to avoid this time consuming procedure, two short synthetic peptides were designed composed of amino acids which flank the splice site present in the MUC1/Y protein. The MUC1/Y peptides used as targets for selection consist of 10 and 20 amino acids, referred to as 10mer peptide and 20 mer peptide respectively. The 10 mer peptide sequence (TEKNAFNSSL) was selected on the basis of 5 amino acids before the splice site and 5 amino acids after the splice site which distinguishes the MUC1/Y from the MUC1 extracellular domain protein sequence. The 20 mer peptide sequence (SVPSS**TEKNAFNSSL**EDPST) is an extension to the 10 mer peptide of 5 additional amino acids either side, hence, 10 amino acids before the splice site and 10 amino acids after the splice site. There are two splicing sites: splice site 1 occurs between amino acids alanine (A) and valine (V) and the splice site 2 occurs between amino acids glutamine (Q) and phenylalanine (F) (figure 3.0). The section between the two splicing sites has the amino acids which constitute the MUC1 tandem repeat array and its flanking sequences. Removal of this central section will result in formation of the MUC1/Y splice variant with a single splice site between the amino acid alanine (A) and phenylalanine (F) [21].

The purpose of the MUC1/Y peptides is to mimic the splice site present in the MUC1/Y protein, which is a specific portion of the amino acid sequence that differentiates the MUC1/Y isoform from the MUC1 protein and its other splice variants. Furthermore, the two peptide sequences used as targets in the selection process will ascertain whether the additional amino acids present in the 20mer peptide influence the binding with the aptamers in comparison to the 10mer peptide. Hence, the comparison of the two resulting aptamer pools will determine any sequence homology within each pool and between the 10mer and 20mer aptamer pools.

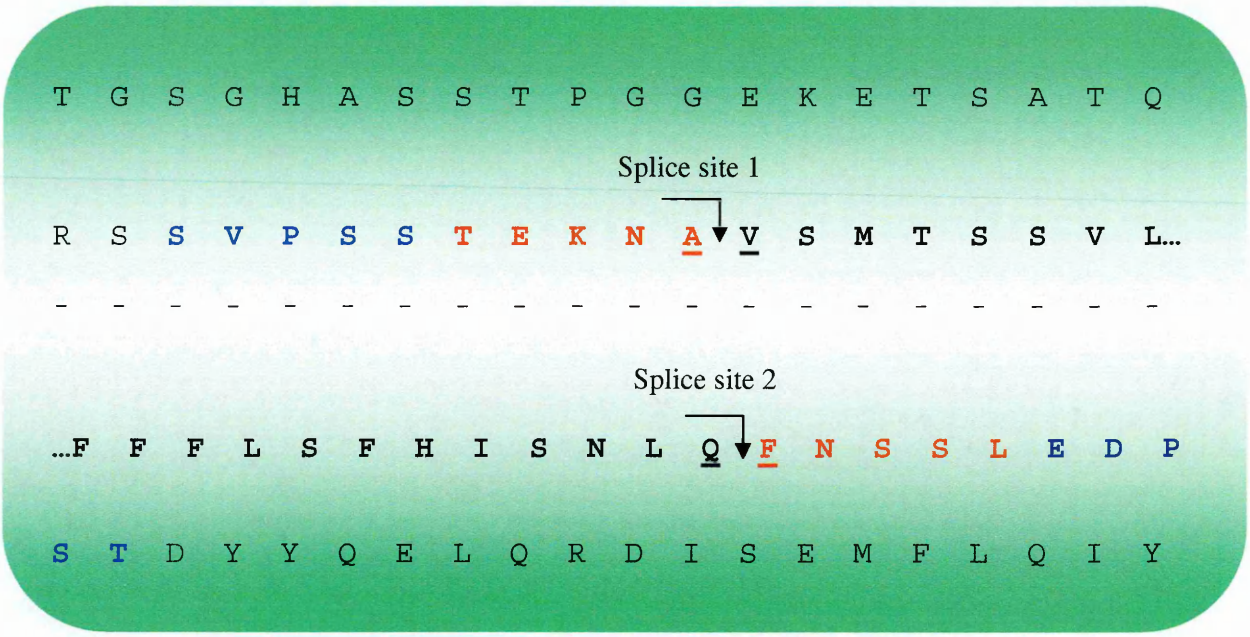


Figure. 3.0 Partial amino acid sequence of MUC1 and MUC1/Y protein. For simplicity, the majority of the spliced out region consisting of the MUC1 tandem repeat array and its flanking sequence have been removed and this is indicated by the dashed line. The amino acids in red are the 5 amino acid flanking the splice site representing the 10 mer peptide and the 5 amino acids in blue on either side of the 10mer peptide represent the 20 mer peptide. Adapted from ref [21].

3.4 BIOTINYLATION OF THE MUC1/Y PEPTIDES

The SimpLex selection process, on this occasion was performed utilising streptavidin coated tubes for target immobilisation. Thus, to immobilise the peptides to the surface of these tubes and ensure that the peptides remain attached to the surface during the entire selection, it was necessary to biotinylate the peptides. Biotin has an extremely high affinity, a K_D of 4 x 10^{-14} M [218] for the tetrameric streptavidin protein and will form a stable irreversible bond with the protein that is strong yet non-covalent. The bond between streptavidin and biotin typically remains unaffected in high salt concentrations, solvents [218], high temperatures and wide range of pH solutions [219].

The MUC1/Y peptides were labelled with biotin-XX protein using a FluoReporter® Biotin-XX protein labelling kit. Each peptide was dissolved in 0.1M sodium carbonate, pH 8 and incubated with the biotin-XX succinimidyl ester (SE) in DMSO for 2 hours at ambient temperatures. The solution of biotin-XX SE in DMSO was only prepared the moment prior to adding to the peptide solution to prevent the hydrolysis

of SE. Subsequent to incubation, the reaction mixture was purified with the aim to separate the biotinylated peptide from the free N-hydroxy succinimide (NHS) and biotin-XX, using a size exclusion PD-10 column with PBS as the eluant. The first 3 mL of flow through were collected in 1 mL fractions and the final fourth 1 mL was collected in 0.5 mL fractions (washout). Biotinylation of the peptides was carried out according to the manufacture's instructions with the exception of using a 1:2 instead of the recommended 1:10 molar ratio of peptide to biotin-XX SE. A reduced ratio of biotin-XX SE was used to avoid biotin conjugation taking place at more than one primary amine sites present in the peptide sequence. To determine the presence of the biotinylated peptides in the individual fractions collected, they were analysed by UV-Vis spectroscopy using an absorbance wavelength scan ranging from 200 nm to 600 nm (figure 3.1). The 10 mer and 20 mer peptides were detected by the absorbance of the amino acid phenylalanine at 256 – 258 nm. For reference, biotin-XX SE and DMSO were also analysed to also determine their presence in any of the fractions collected. The UV absorbance of the flow through from the biotinylated 10 mer peptide has shown presence of the peptide in fractions 2 and 3. The UV absorbance of the flow through from the biotinylated 20 mer peptide indicates the presence of the peptide mainly in the fractions 2 and 3 and possibly a small amount in fraction 1. However, there is a possibility that these fractions could also contain the presence of the unreacted biotin-XX as the biotin-XX has a similar absorbance wavelength as the peptides.

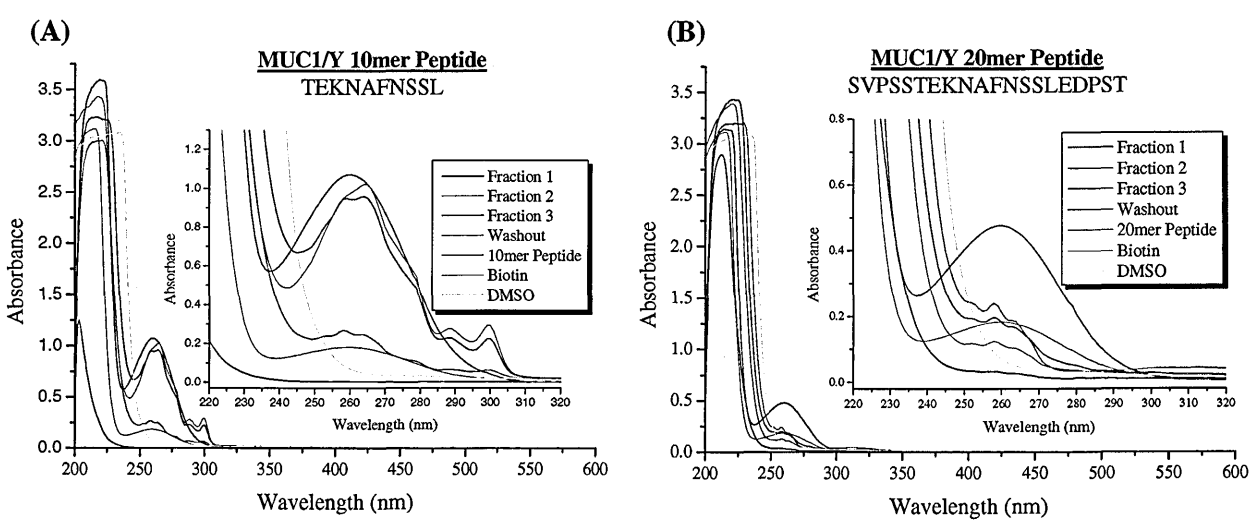


Figure 3.1. UV absorbance spectra of the biotinylated MUC1/Y peptides. The absorbance at 256 – 258 nm was considered for the detection of **(A)** the biotinylated 10 mer peptide and **(B)** the biotinylated 20 mer peptide.

From the UV absorbance spectra, the purity of each fraction cannot be established and neither can the conjugation of the peptides with the biotin-XX SE be determined. Thus, to ascertain whether biotinylation of the peptides has taken place and that the fractions containing the peptides are pure and free of any unreacted biotin-XX, these fractions were further analysed and purified by RP-HPLC. The non-biotinylated 10mer peptide had a retention time (R_t) of 9.34 minutes, which also appeared to contain an impurity that eluted at 15.99 minutes (Figure 3.2). The biotin conjugated 10mer peptide had a delayed R_t of 15.97 minutes. Biotin-XX alone has an R_t of *ca.* 3.5 minutes and the presence of biotin-XX has been observed in both of the biotinylated peptide solutions (fraction 3) which had eluted at 3.48 minutes.

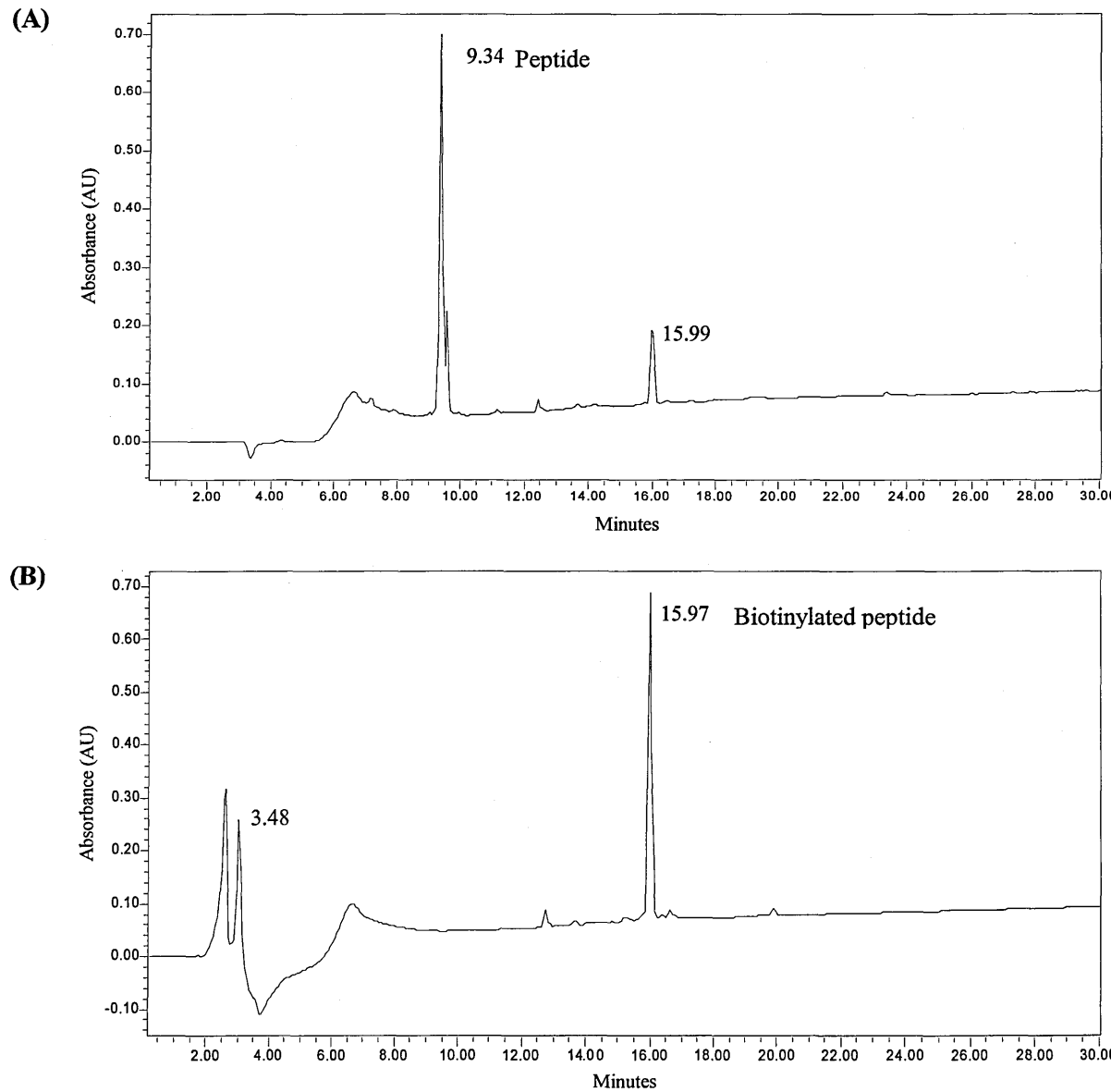


Figure 3.2. HPLC chromatograms of the MUC1/Y 10mer peptide. **(A)** The non-biotinylated 10mer peptide eluted at 9.34 minutes. **(B)** The biotinylated 10mer peptide in fraction 3 eluted at 15.97 minutes.

The non-biotinylated 20mer peptide had a retention time (R_t) of 10.03 minutes, which also appeared to contain an impurity that eluted at 12.89 minutes (Figure 3.3). The biotin conjugated 20mer peptide had a delayed R_t of 15.46 minutes.

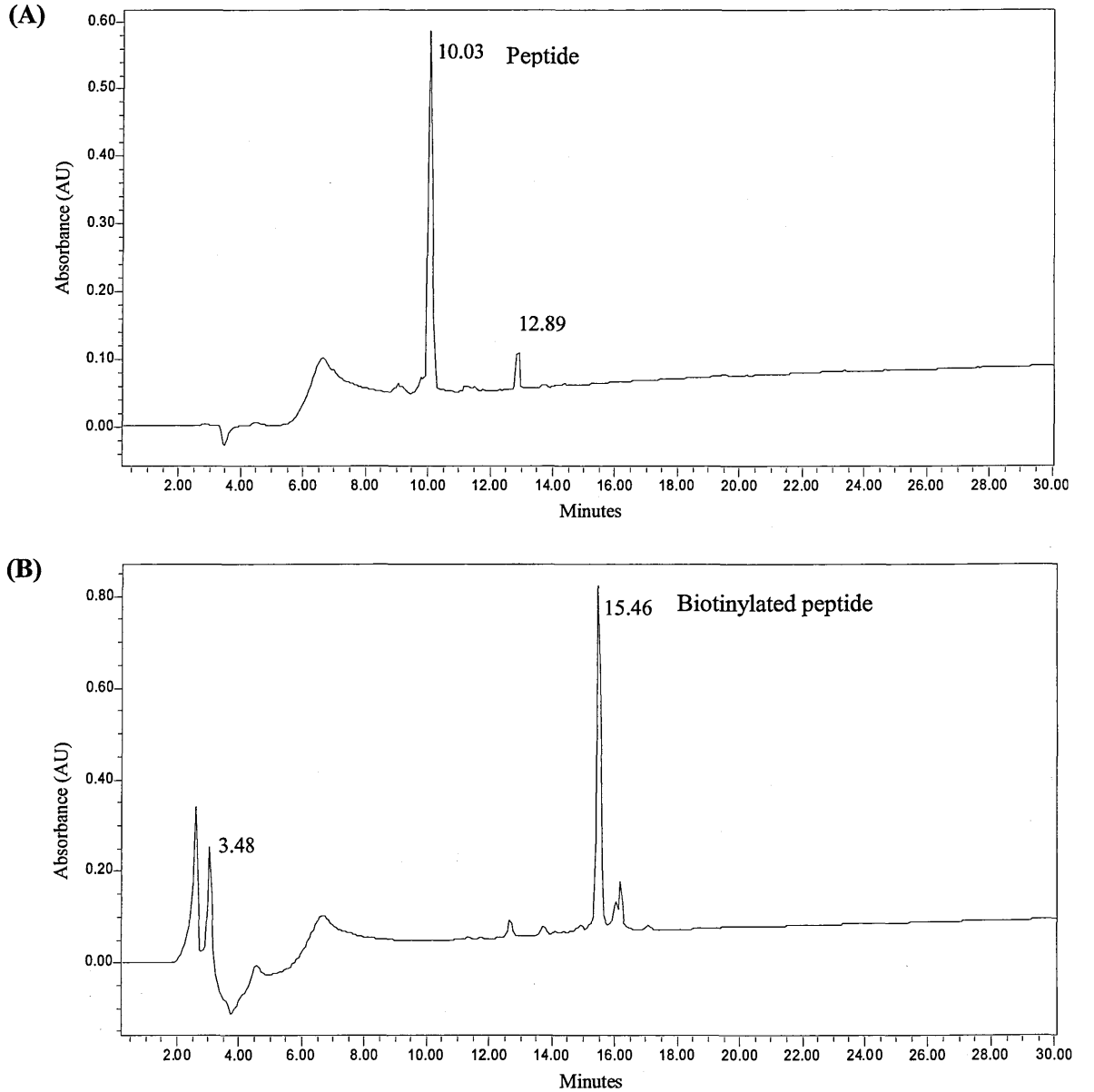


Figure 3.3. HPLC chromatograms of the MUC1/Y 20mer peptide. **(A)** The non-biotinylated 20mer peptide eluted at 10.03 minutes. **(B)** The biotinylated 20mer peptide in fraction 3 eluted at 15.46 minutes.

The HPLC analysis of the 10mer and the 20mer biotinylated peptides have evidently shown that both of the MUC1/Y peptides have been biotinylated, indicated by the increase in retention times in comparison to the retention times of the peptides prior to biotinylation. Furthermore, the biotinylation reaction

appears to have occurred in great efficiency, since unreacted peptide was not detected in all the analysed fractions, which would otherwise be denoted by the retention times corresponding to that of the free peptide. However, analysis of the collected fractions did indicate the presence of the biotin-XX in both of the biotinylated peptide solutions. Hence, all the fractions which contained the biotinylated peptides were purified by RP-HPLC.

3.5 THE COMBINATORIAL LIBRARY & PRIMERS

Each DNA strand in the single stranded combinatorial oligonucleotide library is composed of a central randomised region of 25 nt flanked by hybridising primer sequences. The forward hybridising primer sequence is 23 nt on the 5′ end and the reverse hybridising primer sequence is 24 nt on the 3′ end of the variable region. Thus, each DNA strand is 72 nt long in length (Figure 3.4a). The PCR amplification primers consist of the forward primer, which has a sequence identical to the forward hybridising sequence, and the reverse primer sequence is complementary to the reverse hybridising primer (Figure 3.4b). These primer sequences were used in the selection of aptamers for MUC1 [197].

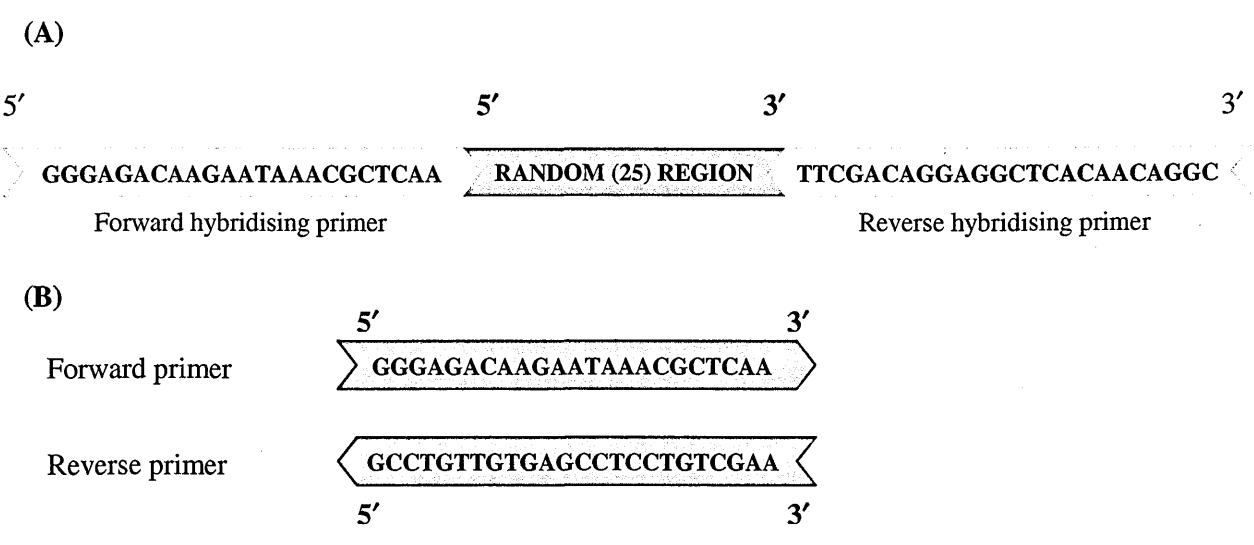


Figure 3.4. The SELEX DNA library and PCR amplification primers. **(A)** Single stranded DNA template entailing hybridising primer sequences flanking a randomised region of 25 nt. **(B)** PCR amplification forward and reverse primers.

The 72 nt long DNA sequence is appropriate in length to generate secondary structures, such as stems and loops, hairpins or G-quadruplexes. In addition, a random sequence of 25 nt gives a molecular diversity of 4^{25} which approximately provides 10^{15} different sequences in the library of single stranded DNA oligonucleotides, hence resulting in numerous different conformations. This vast diversity of structural conformations is likely to provide a pool of DNA molecules with a range of affinities, accompanied by the involvement of the different type of interactions with its target.

3.6 AMPLIFICATION OF THE SELEX LIBRARY

The combinatorial single stranded DNA library was asymmetrically PCR amplified using a forward and reverse primer ratio of 100:4 respectively to generate additional copies of each sequence. The increased number of copies will allow each sequence with its unique folded structure a greater opportunity to compete with the other sequences within the pool for the target. During the PCR amplification, a hot start PCR approach was employed to prevent any incorrect amplification from occurring due to potential mispriming or primer dimerization. The likelihood of mispriming and primer dimerization occurring is typically at low temperatures (ambient temperature) during PCR sample preparation and until the sample reaches the optimal annealing temperature. The hot start PCR method entails withholding an essential reagent for PCR amplification, Taq polymerase, until the reaction mixture has reached a temperature that inhibits primer hybridisation to any non-target sequences. The enzyme Taq polymerase was thus withheld until the reaction mixture reached 95°C (denaturation step) and was added prior to the annealing step at 56°C. Upon annealing of the primers, the temperature is raised to 72°C, optimum for Taq polymerase, whereby enzymatic extension of the primers can take place. This process can aid in increasing the possibility of ensuring the target sequences are amplified correctly.

3.7 IMMOBILISATION OF MUC1/Y PEPTIDES TO THE STREPTAVIDIN COATED TUBES

The biotinylated MUC1/Y peptides were immobilised onto the streptavidin coated tubes according to the manufacturer's protocol in preparation for aptamer selection. Two tubes were immobilised with the 10mer peptide and another two tubes with the 20mer peptide. For each peptide, one tube was utilised in

the salt gradient selection method and the second tube was utilised in the temperature gradient selection. The binding capacity of the streptavidin coated tubes is 15ng/tube for biotin and the immobilisation of the biotinylated peptides was performed in a ratio of 1:2 biotin binding capacity to the biotinylated peptide (30ng). The biotinylated peptides in PBS, pH 7.2 were incubated in the streptavidin coated tubes for 3 minutes at 37°C. Subsequent to incubation, excess biotinylated peptide was removed and the tubes were washed several times with the binding solution (100 mM NaCl, 20 mM KCl, 5mM $MgCl_2$), and finally blot dried for the aptamer selection.

3.8 THE SALT GRADIENT SELECTION METHOD

The amplified single stranded DNA combinatorial library was added to the tubes coated with the pre-immobilised MUC1/Y peptides and incubated for 1 hour at room temperature with gentle shaking. Upon completion of the incubation period, excess aptamer library was removed and the tubes were washed once with binding solution to ensure the non-specific sequences were removed. A gradient elution of the bound sequences was performed by washing the tubes with an elution solution (0.2 M - 1.5 M NaCl, 20 mM KCl, 5mM $MgCl_2$) of increasing concentrations of NaCl starting from 0.2 M to 1.5 M at 0.1 M increments (Figure 3.5). Each elution solution was incubated with the bound sequences for 1 minute at ambient temperature with gentle mixing. Following the completion of the elution step, washings obtained from the NaCl concentrations (1.5M, 1.4 M, 1.3 M, 1.2M) were desalted using microcons (3 kDa MWCO).

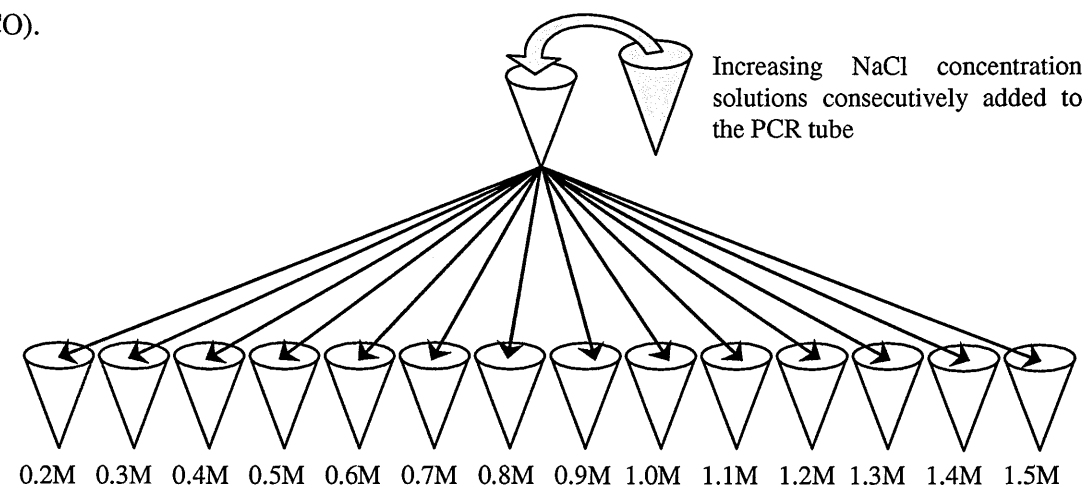


Figure 3.5. The salt gradient elution method. The MUC1/Y peptide bound sequences were eluted with an elution solution of increasing NaCl concentrations, ranging from 0.2 M to 1.5 M in 0.1M increments.

NaCl was used as a chaotropic agent to destabilise the aptamer-peptide complex formed by non-covalent interactions such as hydrogen bonding, Van der Waals or ionic bonding. By employing a NaCl gradient elution system, the bound sequences were separated based on their affinity and specificity for the individual peptides. The weaker binding oligonucleotides expected to dissociate from the peptide at lower NaCl concentrations, and likewise the stronger binders to elute at the higher molarity salt solutions.

The desalted samples obtained from the highest four NaCl concentrations (1.5M, 1.4 M, 1.3 M, and 1.2M) were PCR amplified to generate dsDNA and analysed on a 2% agarose gel to determine the presence of DNA (Figure 3.6). The presence of any oligonucletotides is indicated by observing the DNA at *ca.* 75 bp.

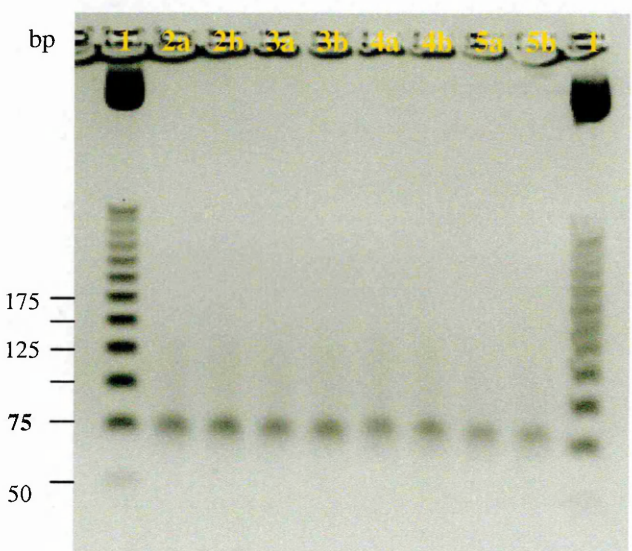


Figure 3.6. 2% agarose gel: Screening for aptamers in varying NaCl concentration samples. (1) 25 bp ladder, **2** – 1.5 M: **(a)** 10mer, **(b)** 20mer, **3** – 1.4 M: **(a)** 10mer, **(b)** 20mer, **4** – 1.3 M: **(a)** 10mer, 25 **(b)** 20mer, **5** – 1.2 M: **(a)** 10mer, **(b)** 20mer.

From the gel, it is evident that there were oligonucleotides present in all four NaCl concentrations eluted from both of the MUC1/Y 10mer and 20mer peptides. At this stage it cannot be ascertained whether the samples contain a pool of oligonucleotides of identical or different sequences. As potential aptamers were detected at the highest NaCl concentration of 1.5 M, this fraction was taken further for cloning and sequencing.

3.9 THE TEMPERATURE GRADIENT SELECTION METHOD

As with the salt gradient elution method, the amplified aptamer library was incubated with the pre-immobilised MUC1/Y 10mer and 20mer peptides in the PCR tubes for 1 hour at room temperature with gentle shaking. Following incubation the excess library was removed and the tubes were washed once with binding solution to ensure the unbound sequences were removed. A temperature elution gradient of the bound sequences was performed by washing the tubes with equilibrated sterile water at increasing temperatures from 25°C to 95°C, increasing by 5°C increments, with gentle mixing for 2 minutes during each step (Figure 3.7). The increase in the temperature allows denaturation of the DNA to take place and thus, the disruption in the interactions involved between the aptamer-peptide complex and of the aptamers' secondary structures can cause the release of the bound sequences from the peptides into the agitating water.

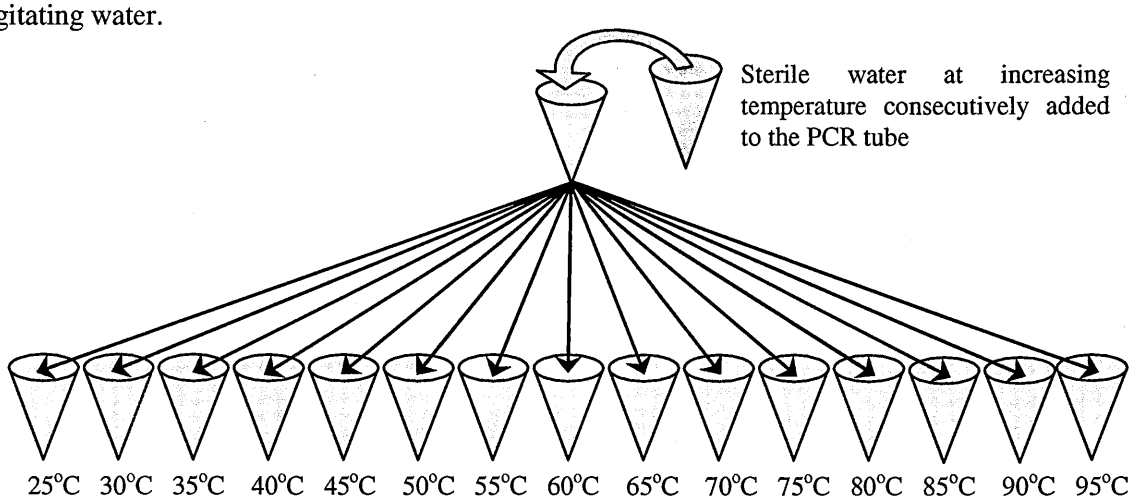


Figure 3.7. The temperature gradient elution method. The MUC1/Y peptide bound aptamers were eluted with equilibrated sterile water at increasing temperature, ranging from 25°C to 95°C in 5°C increments.

Similar to the NaCl gradient elution system, the temperature gradient elution method also dissociates the aptamers from their target peptides based on their affinity and specificity. The less stable binding aptamers were expected to be released from the complex at lower temperatures whilst, the stronger binding sequences were presumed to be eluted at much higher temperatures, as it was expected that strong binding would offer stabilisation of the DNA structure and complex formation. The samples obtained from the highest three temperatures (85°C, 90°C, and 95°C) were PCR amplified to generate

dsDNA and analysed on a 2% agarose gel to determine the presence of aptamer (Figure 3.8). The presence of any oligonucleotides is indicated by observing the DNA at *ca.* 75 bp.

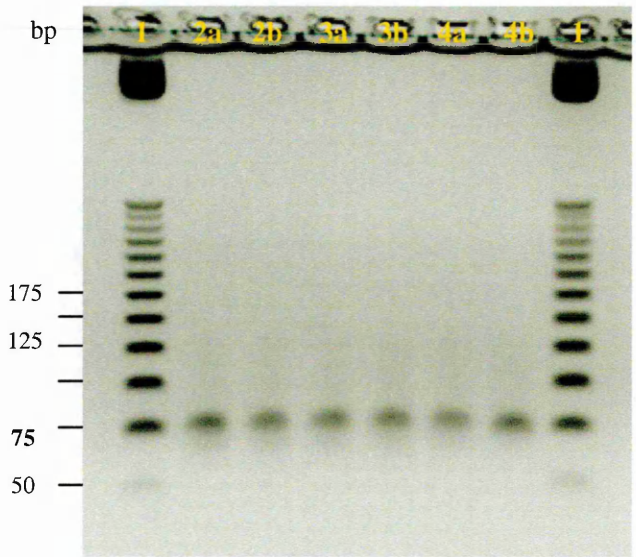


Figure 3.8. 2% agarose gel: Screening for aptamers in varying temperature samples. (1) 25 bp ladder, **2** – 95°C: **(a)** 10mer, **(b)** 20mer, **3** – 90°C: **(a)** 10mer, **(b)** 20mer, **4** – 85°C: **(a)** 10mer, 25 **(b)** 20mer.

The presence of oligonucletotides in all three temperature elution samples obtained from both the MUC1/Y 10mer and 20mer peptides is observed from the gel electrophoresis analysis. The oligonucleotide sample obtained at the highest temperature of 95°C was taken further for cloning and sequencing.

3.10 CLONING & SEQUENCING

The amplified oligonucleotide pools of the NaCl concentration (1.5 M) and temperature (95°C) for both 10mer and 20mer peptides were cloned into the TOPO vectors, using the TOPO TA cloning kit, according to the manufacture's protocol and as described in 2.10. Once the product is cloned into the PCR® 2.1 TOPO vector, this recombinant vector is then transformed into competent E.Coli cells and grown onto LB agar plates overnight at 37°C. Individual colonies of 1.5M and 95°C samples from both peptides grown onto the LB agar plates were swabbed for PCR amplification, as described in 2.11., and analysed on a 2% agarose gel. A successful insert of a 72 bp

oligonucleotide into the vector between the M13 forward and reverse primers will be determined by the presence of a DNA band observed at 272 bp referred to as a positive clone. On the other hand, a negative clone will be detected by a band at 200 bp representing only the M13 forward and reverse primers and the region between the primers within the vector, lacking the 72 bp aptamer insert. Colonies from the 1.5M and 95°C samples for both peptides were analysed until 10 positive clones were obtained from each sample (Figure 3.9).

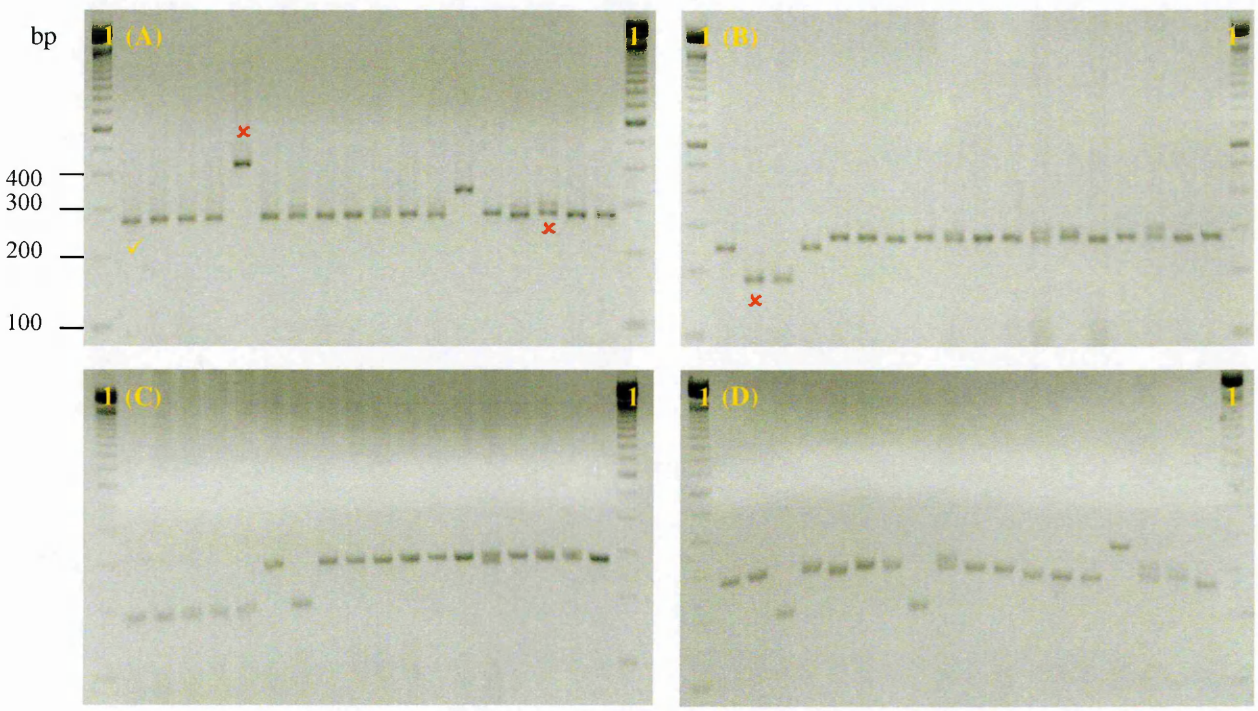


Figure 3.9. 2% agarose gel: Screening for positive clones. **(1)** 100 bp ladder. Colonies from 1.5 M fractions: **(A)** 10mer peptide, **(B)** 20mer peptide. Colonies from 95°C fractions: **(C)** 10mer peptide, **(D)** 20mer peptide. ✓ = indicates an example of a positive clone at 272 bp. ✗ = indicates examples of negative clones, lacking the 72 bp insert within the vector.

Forty randomly selected positive clones (10 each from the 1.5M and 95°C samples for both peptides) were sent for sequencing to Macrogen (South Korea) to identify the 25 nt unknown random region.

3.11 SEQUENCE & M-FOLD STRUCTURAL ANALYSIS

The forty aptamer sequences identified (appendix 1.0) displayed no sequence homology, thus a variety of aptamer sequences were obtained from both selection methods for both of the MUC1/Y 10mer and

20mer peptides. However, selection of aptamers against the MUC1/Y 20mer peptide was repeated (by Dr Huma Khan in the aptamer group at the Open University) with a protocol using the standard SELEX method which employed an affinity chromatography system, entailing several rounds of selection. The aptamer library was incubated with the 20mer peptide in potassium phosphate buffer with salts (100 mM, 20 mM KCl, and 5mM $MgCl_2$), pH 7.2 for 1 hour. Subsequent to incubation, the bound sequences were eluted with increasing salt concentrations from 1 M to 1.5 M in 0.1 M increments and then with a final wash with 3 M NaSCN. The aptamers obtained at 3M NaSCN were subjected to further rounds of selection which was then followed by cloning and sequencing. Oligonucleotide sequences obtained from this SELEX protocol displayed sequence consensus for three different sequences, S11a, S51a and S75a (table 3.0), which were named after the number allocated to the colony sent for sequencing. These three sequences were also present within the set of sequences obtained from the SimpLex system using the salt and temperature gradient elution method. Hence, the sequence homology obtained for these three sequences in the affinity chromatography selection, along with their elution at the highest salt concentration and temperature in the Simplex system, was the rationale behind taking S11, S51 and S75 forward for further investigations.

Name of Sequence	Forward hybridising primer	Central variable region	Reverse hybridising primer
S11a	GGGAGACAAGAATAAACGCTCAA	GGCAACATACTGTAAAGCTCAGGAC	TTCGACAGGAGGCTCACAACAGGC
S51a	GGGAGACAAGAATAAACGCTCAA	AGATAAAAAGGCTGTCTGAAAATT	TTCGACAGGAGGCTCACAACAGGC
S75a	GGGAGACAAGAATAAACGCTCAA	CTGGTATCTATATGAAGGTTGTAGC	TTCGACAGGAGGCTCACAACAGGC

Table 3.0 Aptamer sequences of S11a, S51a and S75a. In blue are the forward and reverse hybridising primer sequences necessary for PCR amplification. In red are the variable (previously unknown) regions identified through sequencing.

From the SimpLex system, sequences S11a and S75a were selected against the 10mer peptide which was eluted at the 1.5 M NaCl concentration, whilst the aptamer S51a was also selected against the 10mer peptide but eluted at 95°C. However, these sequences were obtained under different conditions in the affinity chromatography SELEX method, where they were selected against the 20mer peptide and eluted

with 3 M NaSCN. Thus, this can be an indication that these three oligonucleotides are capable of forming a complex with both of the MUC1/Y peptides. Furthermore, another possible insight to their means of binding to the peptides is that the formation or stabilisation of the potential aptamer secondary structures is not dependent upon the cations present in the solutions during binding and may not be involved in the peptide-aptamer complex formation. This assumption is deduced based on the difference between the solutions in which the aptamer libraries were incubated in with the peptide. In the Simplex system, the combinatorial library was incubated with the peptide in the solution which the library was amplified in, containing PCR buffer and PCR reagents, whilst in the affinity chromatography selection the aptamer library was incubated with the peptide in potassium buffer containing the binding solution salts at pH 7.2.

The predicted secondary DNA structures for the sequences S11a, S51a and S75a were obtained from m-fold web server [220], an algorithm used to predict the secondary structures of DNA and RNA molecules. The predicted structures provide a broad indication of the different type of base parings which can occur to generate the stem and loop structures that could be play an important role in the binding to the MUC1/Y peptides. However, the full length of the aptamers of 72 nt long is not a feasible length for commercial purposes whereby the aptamer yield of such a length is considerably reduced upon synthesis and purification. Also the synthesis of aptamers greater than 45 nt can be expensive and thus non-viable for therapeutic scale-up applications. Therefore, to reduce costs, sequences were truncated for further investigation along with the full length oligonucleotides for comparison. The 72mer sequences were truncated by two alternative means. One was to remove the hybridising primer sequences, leaving only the 25 nt variable region, and these were referred to as S11b, S51b and S75b (table 3.1). The second method was based on their predicted secondary structures obtained from the m-fold structure prediction programme. Sequence truncation by this method was based on removing the bases which were not involved in the formation of the largest stem-loop structure and only incorporating the nucleotides of the secondary structure motif. These sequences were referred to as S11c, S51c and S75c (table 3.1). The predicted secondary structures of the c series of oligonucleotide sequences have shown that the base pairing and structure formation occurs between the reverse hybridising primer and the random variable

region. The forward hybridising primer is not involved in the secondary structure formation. Furthermore, the S75c sequence largely consists of the bases present in the reverse hybridising primer sequences rather the bases of the random variable region in comparison to the S11c and S51c sequences.

The aptamer 'b' series			
Name of Sequence	**Forward hybridising primer**	**Central variable region**	**Reverse hybridising primer**
S11b	GGCAACATACTGTAAAGCTCAGGAC		
S51b	AGATAAAAAGGCTGTCTGAAAATT		
S75b	CTGGTATCTATATGAAGGTTGTAGC		
The aptamer 'c' series			
Name of Sequence	**Forward hybridising primer**	**Central variable region**	**Reverse hybridising primer**
S11c		ACTGTAAAGCTCAGGAC	TTCGACAGG
S51c		GGCTGTCTGAAAATT	TTCGACAGGA
S75c		GGTTGTAGC	TTCGACAGGAGGCTCACAACA

Table 3.1 Aptamer sequences for the 'b' and 'c' series. The b series entails the aptamer sequences which have had the hybridising primers removed and only contain the central variable region. The c series are the aptamer sequences which were truncated based on the predicted secondary structures.

There are several structural possibilities that each aptamer can adopt which is presented by the m-fold predicting programme and the predicted secondary structures of each aptamer and its truncated versions are shown in figures 3.9–3.11. The structures of the S11 oligonucleotide series (Figure 3.10) display a combination of a short hairpin motif and the long hairpin containing a bulge. The full length S11a aptamer contains two hairpins, whilst the truncated versions of the potential aptamer contain a single stem-loop region with the formation of an open structure for the shorter hairpin motifs. The structures of the S51 series (Figure 3.11) are very similar to that of S11 sequences. However, S51b of the central variable region forms less open structures and produces a dumbbell like structure. The S75 series exhibit

structures (Figure 3.12) which also contain stem and loop regions. However, the aptamer S75a forms a long tight hairpin structure and two of the predicted structures show greater base pairing tendency.

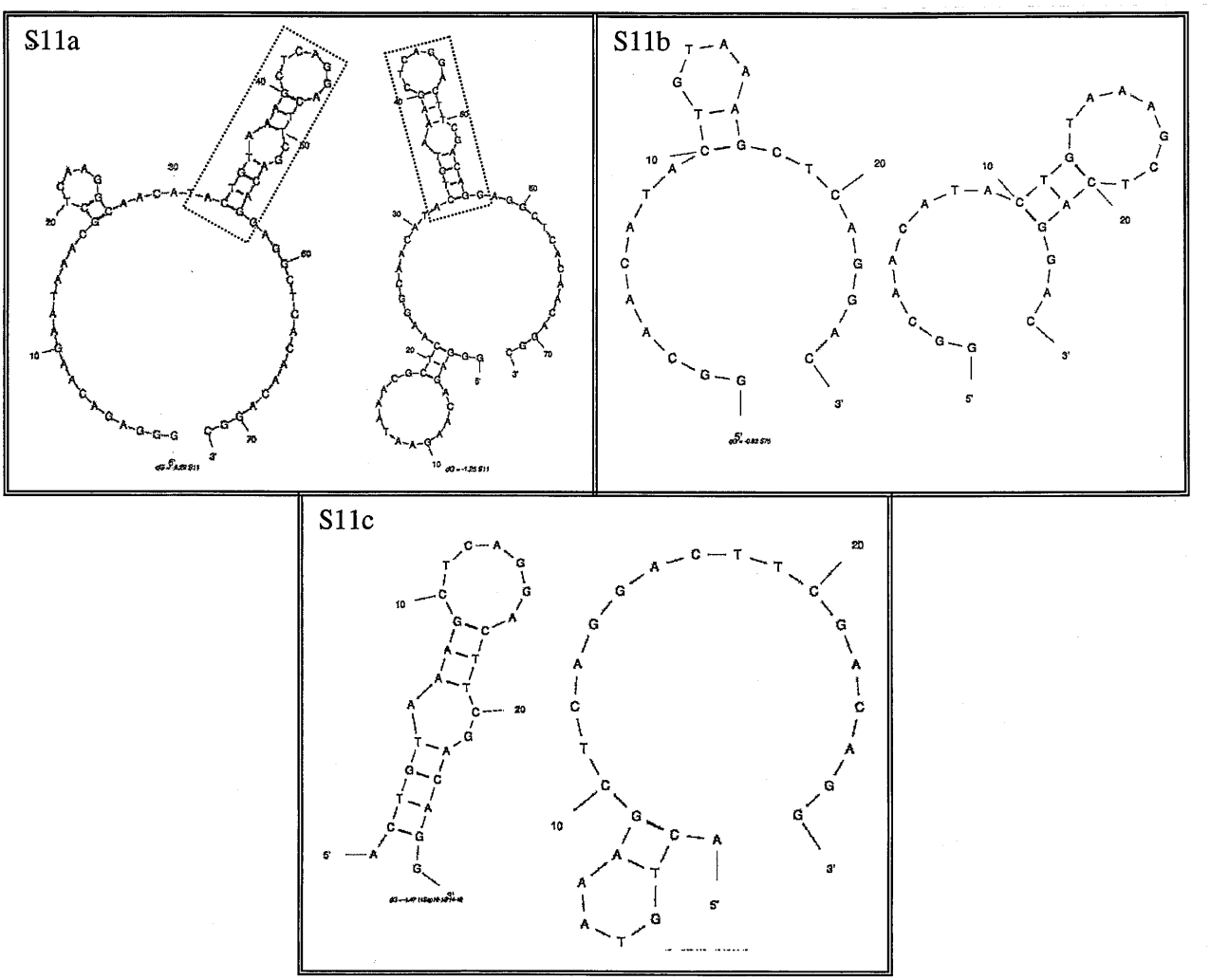


Figure 3.10 The predicted secondary structures of S11 and the truncated versions. S11a of 72 nt (GGGAGACAAGAATAAACGCTCAAGGCAACATACTGTAAAGCTCAGGACTTCGACAGGAGGCTCACAACAGGC) is the full length aptamer. S11b of 25 nt (GGCAACATACTGTAAAGCTCAGGAC) is the variable region. S11c of 26 nt (ACTGTAAAGCTCAGGACTTCGACAGG) is the stem-loop region obtained from the predicted structure of S11a highlighted by orange dotted box. The structure was predicted at 37°C with 100 mM and 5 mM $MgCl_2$.

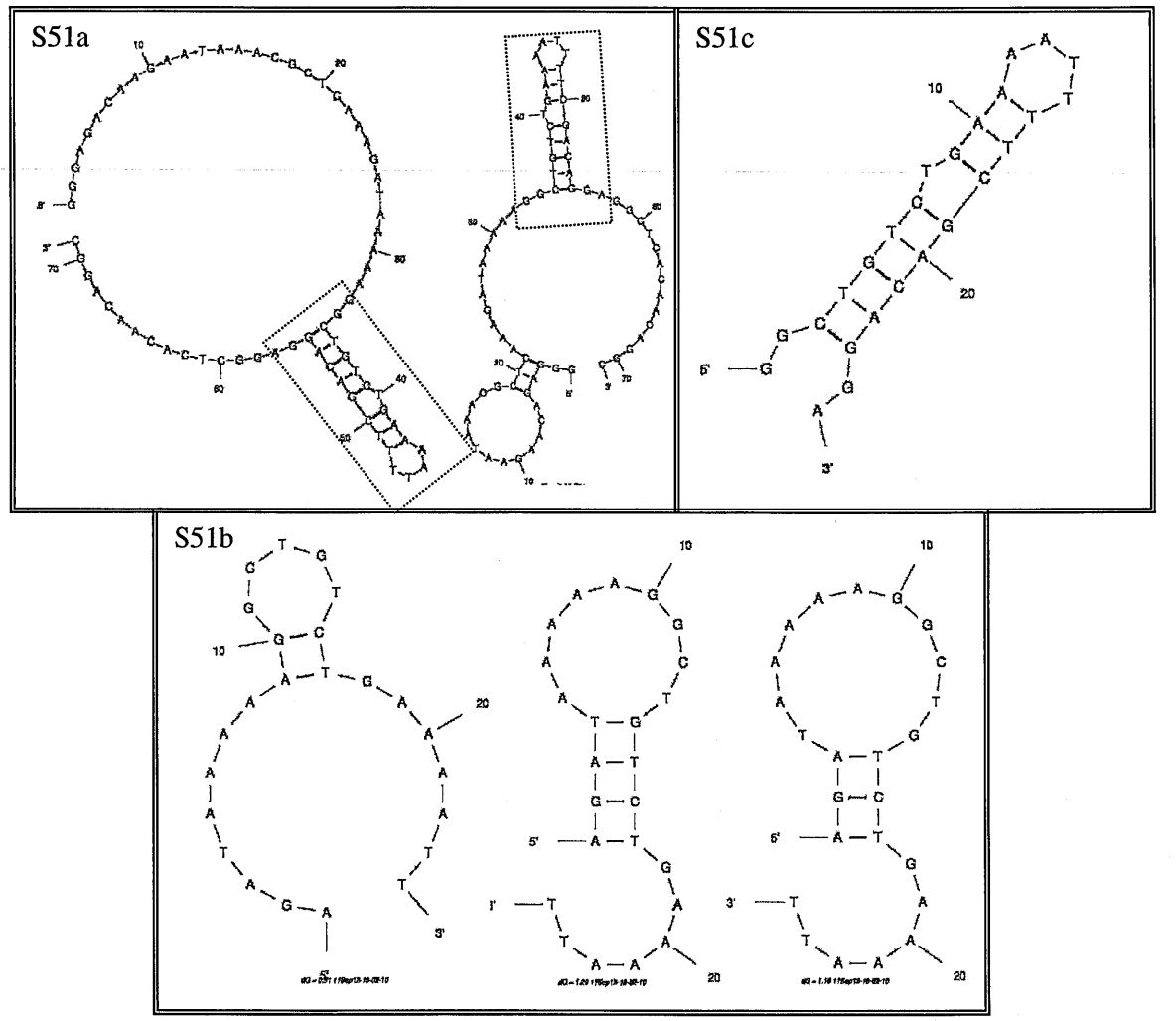


Figure 3.11. The predicted secondary structures of S51 and the truncated versions. S51a of 71 nt (GGGAGACAAGAATAAACGCTCAAAGATAAAAAGGCTGTCTGAAAATTTTCGACAGGAGGCTCACAACAGGC) is the full length aptamer. S51b of 24 nt (AGATAAAAAGGCTGTCTGAAAATT) is the variable region. S51c of 25 nt (GGCTGTCTGAAAATTTTCGACAGGA) is the stem-loop region obtained from the predicted structure of S51a highlighted by orange dotted box. The structure was predicted at 37°C with 100 mM and 5 mM $MgCl_2$.

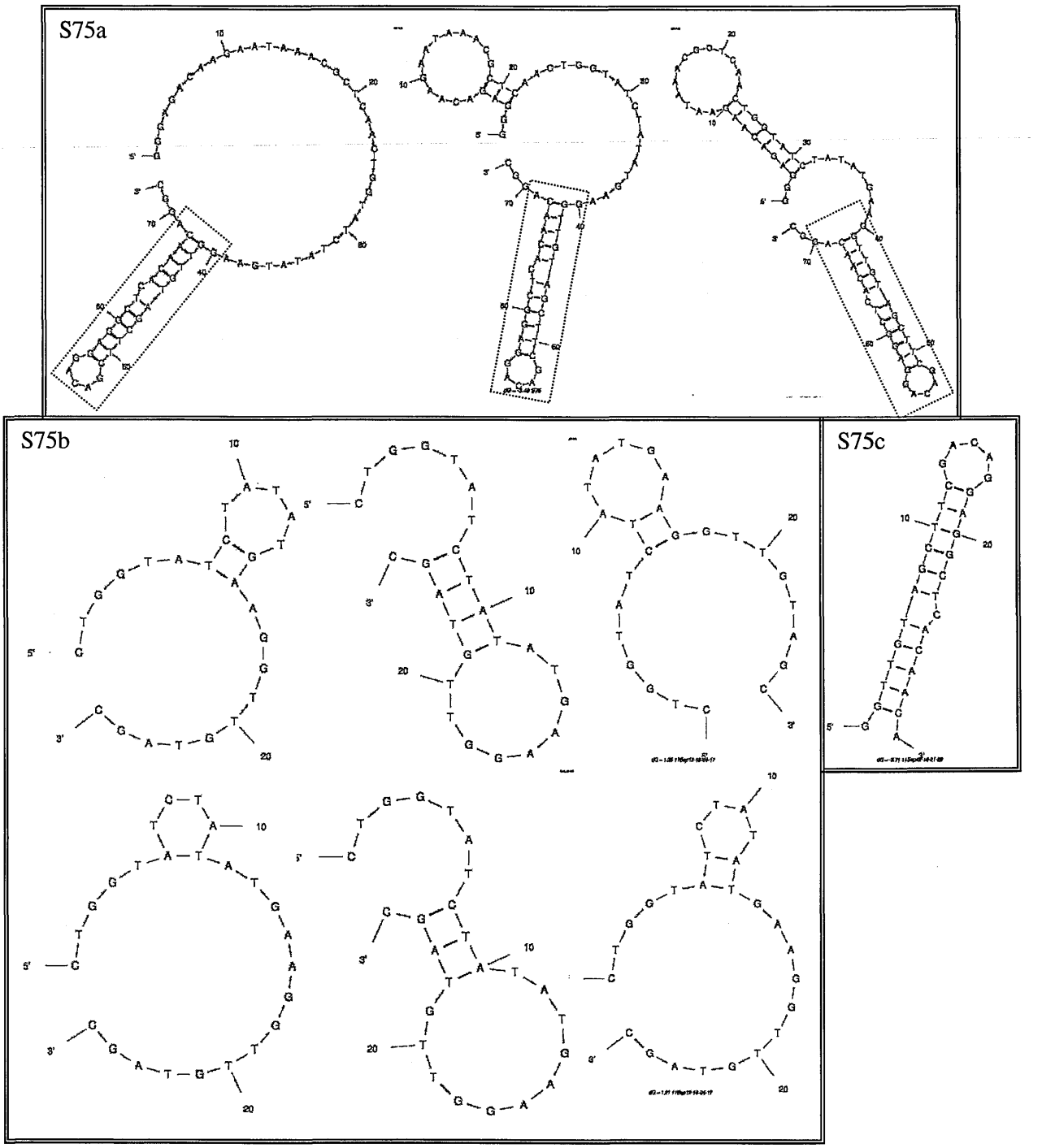


Figure 3.12. The predicted secondary structures of S75 and the truncated versions. S75a of 72 nt (GGGAGACAAGAATAAACGCTCAACTGGTATCTATATGAAGGTTGTAGCTTCGACAGGAGGCTCACAACAGGC) is the full length aptamer. S75b of 25 nt (CTGGTATCTATATGAAGGTTGTAGC) is the variable region. S75c of 30 nt (GGTTGTAGCTTCGACAGGAGGCTCACAACA) is the stem-loop region obtained from the predicted structure of S75a highlighted by orange dotted box. The structure was predicted at 37°C with 100 mM and 5 mM $MgCl_2$.

Thus, a total of nine aptamer sequences were taken forward for investigating further of which the 3 full length aptamer sequences (S11a, S51a and S75a) were the original sequences obtained from the selection. The b series (S11b, S51b and S75b) are the central random region sequences and the c series S11c, S51c and S75c) are the stem-loop regions from the predicted secondary structures.

3.12 CONCLUSIONS

Taking into consideration the traditional SELEX system, involving numerous rounds of selection, and many of the non-conventional aptamer selection methods mentioned previously, an alternative approach for generating aptamers to potentially any target was considered. SimpLex, a one round selection method was developed to obtain aptamers against the MUC1/Y peptides, 10mer and 20mer, possessing high affinity and specificity for their target. In brief, the biotinylated peptides were immobilised to the streptavidin coated PCR tubes, to which the single stranded combinatorial library was added and incubated with the target peptides. The bound sequences were separated according to their binding strength to the target peptides via either a salt gradient or a heat gradient elution. Of the final forty sequences obtained from both elution methods against the 10mer and 20mer peptides, only three sequences (S11a, S51a, S75a) were considered for further studies. The rationale to consider these three sequences in further experimental research was based on the appearance of these sequences from the repeat selection employing the affinity chromatography method against the 20mer peptide. Additionally, another six sequences were designed from the parent sequences, which consisted of two different constructs. The first construct involves the removal of the hybridising primer sequences leaving only the central variable region (S11b, S51b, and S75b) and the second construct entails the sequence forming the stem-loop obtained from the predicted secondary structures.

From the forty sequences that were obtained, sequence homology was not observed between the sequences from either elution methods and amongst both peptides. The lack of sequence consensus could be the outcome of either or a combination of two possible reasons: first, that only a small number of clones for each elution method and peptide were sequenced and second, that the elution methods may not have been stringent enough. The analysis of only a small number of clones could imply that the

probability of acquiring sequence homology is significantly reduced as opposed to screening a hundred or more clones. Moreover, the stringency of the eluting system could have been increased by subjecting the aptamers obtained from the 1.5 M NaCl concentration and 95°C temperature fractions to another round of selection. In addition, to further refine the selection process, the final pool of aptamers from both elution systems could be subjected to negative and counter selections. For a negative selection, the final pool of aptamer could be incubated with only the streptavidin coated tubes, thus removing any potentially non-peptide specific aptamers. The counter selection could be performed by incubating the aptamer pool with the MUC1 peptides [197] which would remove any sequences which may have affinity for a non-selected target, therefore enhancing the specificity of the aptamers obtained for MUC1/Y. Although, potential aptamers S11b, S51b, and S75b were all generated against the 10mer peptide in the SimpLex system and not present in the pool of prospective aptamers obtained for the 20mer peptide, this does not suggest that these three lead sequences do not bind to the 20mer peptide. In fact, the contrary was affirmed as in the affinity chromatography SELEX selection these three lead sequences were generated against the 20mer peptide and eluted from the column at the higher salt concentration of 3 M instead of the 1.5 M. Furthermore, the utilisation of two different selection protocols exemplifies that the three sequences do not possess any affinity for the medium in which the selection took place. For example, in the case of the SimpLex system, the support medium is the streptavidin which coats the PCR tube whilst in the affinity chromatography system an activated sepharose resin was employed. Furthermore, the binding of these three aptamers is not dependent upon the cationic salts in the buffer as in both selection methods the combinatorial DNA library was incubated with the peptides in different cationic salt solutions.

The SimpLex, one round selection system, has shown to have the potential to generate aptamers against a known target in a simple and cost effective manner. The aptamer generating protocol can be replicated on a lab bench without the requirement of any sophisticated and expensive instruments. The key advantages of this protocol are that it is extremely fast, not practically laborious and can be performed using small quantities of reagents. This selection process can be completed in two hours from immobilising the target, incubating the combinatorial library to eluting the bound aptamers. Even if the final pool of

aptamers obtained at the highest NaCl concentration and temperature is subjected to a second round of selection, this can still be completed within a day. Additionally, provided that the target in question can be immobilised on streptavidin coated PCR tubes, there are no other limitations with respect to the size of the molecule which can be used as a target.

CHAPTER FOUR

FUNCTIONALISATION OF MUC1/Y APTAMERS

4.0 INTRODUCTION

Aptamers present remarkable potential as therapeutic agents. However, they cannot be directly taken into *in vivo* studies prior to any modification due to their poor pharmacokinetics properties and sensitivity to nuclease degradation. Unmodified aptamers are subjected to rapid renal clearance, thus preventing the aptamer from reaching the biomarker to result in an adequate therapeutic response. This limitation can be addressed through a chemical modification referred to as pegylation, which typically entails conjugating the aptamer to a high molecular weight polymer, polyethylene glycol (PEG), at both or either of its terminal ends. Pegylation has been shown to alter the pharmacokinetics of aptamers by increasing their blood circulating half-life from minutes or hours to days [221]. This allows the aptamer sufficient time to reach the target of interest and bring about the desired effect. The molecular weight cut off for glomerular filtration is noted to be between 30 kDa and 50 kDa [222], whereby urinary excretion of pegylated molecules appears to be directly proportional to the mass of PEG; urinary clearance reduces as the molecular weight of the PEG increases. On the other hand, liver clearance has an inverse response, as the molecular weight of PEG increases so does liver clearance. However, this effect has been reported to take place only when the minimum size of 50 kDa of the molecule is exceeded [223].

The covalent attachment of PEG to biological molecules was initially introduced to increase the molecular weight of peptides and proteins, amongst many other advantages, which was then later applied to aptamers. The most commonly noted methods described in literature for pegylation involve reacting a functional moiety of PEG with either an amine or a thiol group present in proteins and peptides. Activated functional groups of PEG which can covalently form an amine linkage with an amino group include active ester, active carbonates, aldehyde or tresylate. Thiol group present on the amino acid cysteine can be attached to PEG via thiol reactive derivatives of PEG such as maleimide, vinyl sulfones, iodoacetamide and orthopyridyl disulfide [193,196,224]. Hence, aptamers can be modified to possess the functional groups present in proteins or peptides to be able to react with the activated PEG.

The truncated MUC1/Y aptamers of series b and c (S11b, S51b, S75b and S11c, S51c, S75c, respectively) have been pegylated as the full length MUC1/Y aptamers of 72 nt would not be appropriate

or for therapeutic and commercial applications as explained previously. Several different parameters for aptamer conjugation had been explored before the optimised condition was obtained to provide a high yielding aptamer pegylation reaction.

4.1 POLYETHYLENE GLYCOL CONJUGATED APTAMERS

Numerous aptamers have been conjugated to PEG of different molecular weights, structures and functional groups to alter the pharmacokinetics of the unmodified aptamers. A pegylated aptamer known as Pegaptanib is currently in the market for the cure of age-related macular degeneration. Pegaptanib is a 27 nt long RNA aptamer generated against the $VEGF_{165}$ protein which has an amino modification on the 5′ end and is conjugated to a 40 kDa NHS PEG to reach an overall molecular weight of *ca.* 48 kDa. Due to the conjugation to PEG, the aptamer showed no degradation in plasma at ambient temperature for longer than 18 hours, however a four fold decrease in its binding affinity to the target $VEGF_{165}$ was also detected. Although a reduced binding was observed as a result of pegylating the aptamer, the inhibition of vascular permeability stimulated by VEGF was improved to 83%. It has been suggested that this increase in the inhibitory activity by Pegaptanib is the potential effect of a decrease aptamer rate of diffusion due to the increase in the molecular weight from the conjugation to a 40 kDa PEG polymer [66,157].

An anti-PDGF DNA aptamer, ARC127, used to reduce interstitial tumour pressure for promoting greater uptake chemotherapeutic agents into the tumours has also been conjugated to a 40 kDa PEG polymer. Also in this case, the 29 nt long aptamer with a primary amine group modification at the 5′ end was reacted with the NHS activated PEG ester, increasing the overall molecular weight of the aptamer to *ca.* 49 kDa. The pegylated ARC127 showed an increase in the *in vivo* aptamer half-life from a few minutes to 8 hours when administered intravenously. Furthermore, the pegylated aptamer was able to reduce tumour cell proliferation by 95% over a period of 9 days and inhibited the process induced by the PDGF [170,225].

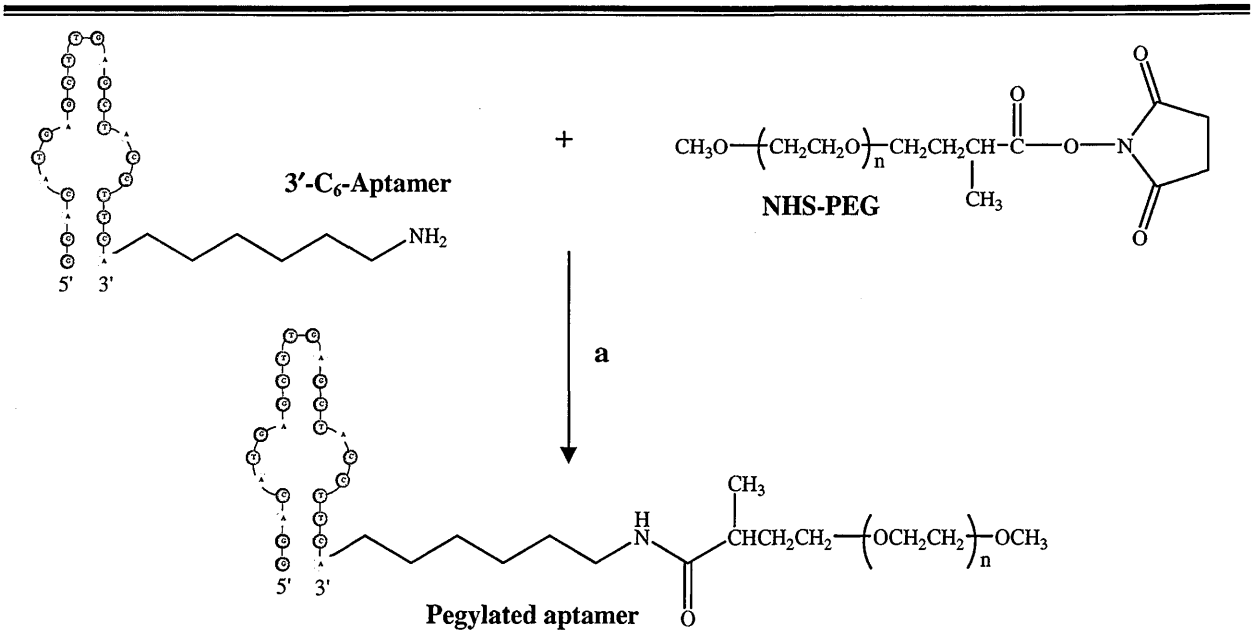
Another therapeutic aptamer which has been conjugated to PEG is the 15 nt long DNA anti-thrombin aptamer employed as an anticoagulant. The anti-thrombin aptamer was attached to a 40 kDa PEG in a different manner, where the aptamer had two C_6 amino linkers at its 5′ end that were reacted with two linear 20 kDa NHS activated PEGs, resulting in a branched type structure. The final molecular weight of the pegylated aptamer was increased to *ca.* 44 kDa. The anti-thrombin aptamer prior to modification had a whole blood clotting time (WBCT) of 382 seconds which was increased to 456 seconds once pegylated [170,226].

The ARC83 aptamer is a 32 nt long scrambled form of the aptamer generated against the transforming growth factor-β2 (TGF-β2). ARC83 having an amino functionality at the 5′ terminus was attached to two different molecular weight PEG polymers 20 kDa and 40 kDa via the activated succinimidyl ester moiety of PEG to form ARC120 (30 kDa) and ARC122 (50 kDa) respectively. The plasma pharmacokinetics in *in vivo* rat models revealed the un-pegylated aptamer had a half-life ($t_{1/2}$) of 5 hours with a mean residence time (MRT) of 1.7 hours. ARC120 showed a higher $t_{1/2}$ of 7 hours with a MRT of 8 hours while ARC122 had a $t_{1/2}$ of 12 hours with a MRT of 16 hours. Both pegylated aptamers have demonstrated significantly slower clearance from the circulatory volume, further enhancing the distribution of the aptamer to the tissues. Moreover, the 40 kDa pegylated aptamer demonstrated an MRT *ca.* 10 fold higher than the unmodified aptamer [221,227].

Anti-human L-selectin RNA and DNA aptamers have also been conjugated to 20 kDa and 40 kDa PEG in the same way described previously for the ARC83 aptamer. The final molecular weights of the aptamers were 30 kDa and 50 kDa when they were pegylated to the 20 kDa and 40 kDa PEG respectively. Studies of the pegylated aptamers revealed that the MRT of the DNA aptamers ≤ 30 kDa varied with the different overall molecular weights, but the MRT of the conjugated aptamers ≥ 30 kDa remained constant. This indicated that renal filtration was the most important rate determining step for low molecular weight DNA aptamers, though this was not the case for the DNA aptamers above the molecular weight of 30 kDa [65].

4.2 PEGYLATION OF THE MUC1/Y APTAMERS

For aptamers to become viable as therapeutic agents it is important that their molecular weight is increased above 30 kDa. To achieve this, in literature, 20 kDa and 40 kDa PEG have been employed, increasing the overall molecular weight of the pegylated aptamers to approximately 30 kDa and 50 kDa respectively. For the pegylation of the truncated MUC1/Y aptamers (S11b, S51, S75b, S11c, S51c, and S75c) with a molecular weight of 7 – 9 kDa, a 32 kDa linear PEG was employed, increasing the molecular weight of the pegylated aptamers to *ca.* 40 kDa. The increase in overall molecular weight of the truncated aptamers is likely to aid in the prevention of rapid renal clearance, as it takes them above the effective cut-off for the retardation in glomerular filtration ranging between 40 – 60 kDa [194,228]. The method typically described in literature for pegylation involving amine coupling was used for conjugating the MUC1/Y aptamers to PEG. The C_6 amino linker is attached to the phosphate group in the 3′ position of the ribose at the beginning of the solid phase synthesis of the oligonucleotides. The amino group can then react with the activated NHS-ester of 32 kDa linear PEG forming a stable covalent bond (scheme 4.0).



Scheme 4.0. Aptamer pegylation. The 3′-amino modified aptamers were reacted with NHS-activated PEG. **(a)** 53.4 mM sodium carbonate buffer, pH 10, 60°C, 1 hour. n = 736 ethylene oxide monomers.

Initially, the reaction was performed following a combination of protocols described in literature. The aptamer was reacted with a 3 molar excess of the NHS activated PEG in 100 mM sodium carbonate buffer (pH 9.3) for 1 hour at ambient temperature [221,229] but no product was obtained based on the analysis performed by agarose gel electrophoresis (figure 4.0, lane 4). The presence of a pegylated aptamer was expected to be indicated by an aptamer band (stained with ethidium bromide (EtBr)) significantly retarded within the gel compared to the unmodified aptamer (lane 5) which migrates at a faster rate and is observed at 25 bp. Conditions such as pH, temperature and aptamer to PEG ratios were altered with the aim to achieve a high yielding pegylation reaction. Using 53.4 mM sodium carbonate buffer (pH 10), 1: 20 molar ratio of aptamer to NHS-PEG and performing the reaction at 37°C for 1 hour resulted in aptamer conjugation even though significant unreacted aptamer was also observed (figure 4.0_Lane 3). The pegylation reaction was also analysed by agarose gel electrophoresis where a new band close to the well in lane 3 was observed which do not correspond with the starting reagents. This may indicate the formation of a potential aptamer-PEG conjugate.

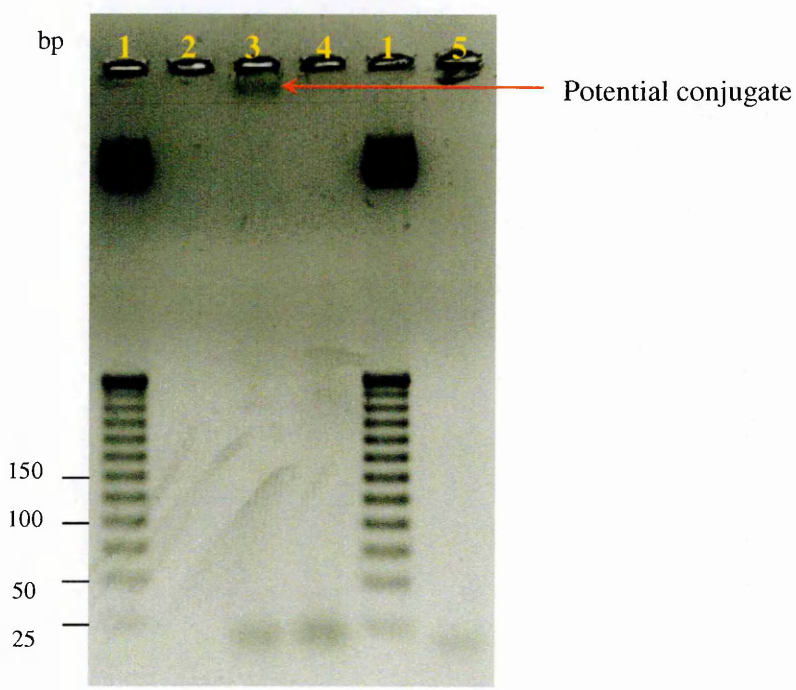


Figure 4.0. 2% agarose gel: Analysis of aptamer-PEG reaction mixtures. (**1**) 25 bp ladder, (**2**) free PEG, (**3**) 1: 20 aptamer: PEG molar ratio, (**4**) 1: 3 aptamer: PEG molar ratio, (**5**) free aptamer (25 nt).

To optimise the reaction further a 1: 40 aptamer: NHS-PEG molar ratio was used and analysed by gel electrophoresis (figure 4.1). Gel A, stained with EtBr showed the potential aptamer conjugate from the 1: 40 molar ratio reaction mixture in lane 4 close to the well, along with unreacted aptamer corresponding to the free aptamer in lane 5. Staining from this indicated that majority of the aptamer had reacted with the activated PEG, as the unreacted aptamer had not stained with the same intensity as the free aptamer. Gel B was stained with a barium chloride iodine (BaClI) solution to detect the presence of PEG as described in literature [230-232]. This form of staining is based on a complex formation between the barium ions and PEG. When iodine solution is added to this complex, the complex becomes insoluble resulting in a visible rusty coloured stain in the gel. Gel B, showed that the free PEG had migrated towards the cathode in lane 4 which corresponds with the free PEG in lane 2. Although the free PEG had stained quite significantly in lane 4, the conjugated aptamer had not stained with the BaClI solution as expected.

The concentration of the potential conjugate may not have been sufficiently high for the formation of a detectable barium iodide-PEG complex. Therefore the 1: 40 molar ratio reaction mixture was also analysed by 7 % native PAGE. Gel C, after EtBr staining in lane 2 is the free aptamer and lane 1 is the pegylation mixture where the potential conjugated aptamer is detected close to the well and the unreacted aptamer has migrated with the 25 bp marker. In this case, the band corresponding to the unreacted had the same intensity of staining as the free aptamer. Once treated with BaClI solution, the band of the potential conjugate was clearly visible as a rusty coloured stained in gel D.

To ascertain whether the reaction has led to aptamer covalently bound to PEG and not only an entrapment of the aptamer within the large PEG polymer a control reaction was performed. The control reaction involved repeating the previous reaction (1: 40 aptamer: NHS-PEG molar ratio). Prior to the addition of PEG to the aptamer, NHS ester moiety of PEG was hydrolysed to release the unreactive carboxylic group by incubating the NHS- PEG in 53.4 mM sodium carbonate buffer (pH 10) at 37°C with gentle mixing for 2 hours. Subsequently, the hydrolysed PEG solution was added to the aptamer and the mix was incubated at 37°C for a further 1 hour. The control reaction was analysed by agarose gel electrophoresis (figure 4.1). In gels A and B, lane 3, only the band of free aptamer visible at 25 bp and

the free PEG were visible indicating that the conjugation had not taken place. Therefore, indicating the new band in lane 4 is not present due to entrapment of aptamer in the large PEG molecule and is likely to be the pegylated aptamer.

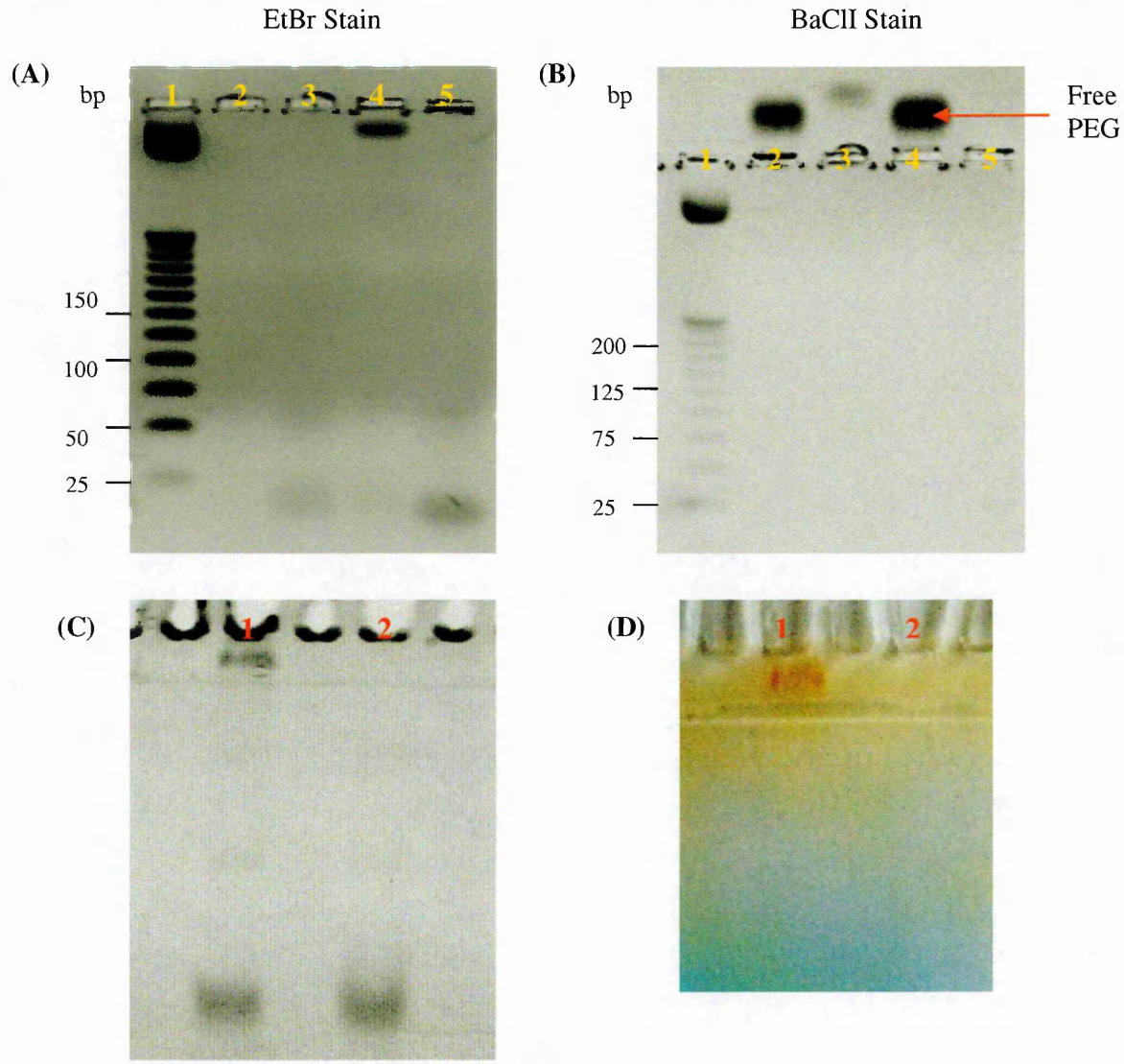


Figure 4.1. 2% agarose gel: Analysis of aptamer-PEG reaction mixtures. (A-B): (1) 25 bp ladder, **(2)** free PEG, **(3)** control reaction, **(4)** 1: 40 aptamer: PEG molar ratio, **(5)** free aptamer (25 nt). **7% native gel: (C-D): (1)** 1: 40 aptamer: PEG molar ratio, **(2)** free aptamer (25 nt).

The control and pegylation (1: 40 aptamer: PEG molar ratio) reaction mixtures were also analysed by anion exchange HPLC (figure 4.2). The chromatogram of the control reaction (figure 4.2A) showed that the hydrolysed PEG and NHS have an R_t of 2.29 minutes and the unreacted aptamer elutes at 16.68 minutes. The chromatogram of the pegylation reaction mix (figure 4.2B) showed new peak having an R_t of 12.26 minutes with a yield of 13 % quantified from peak area.

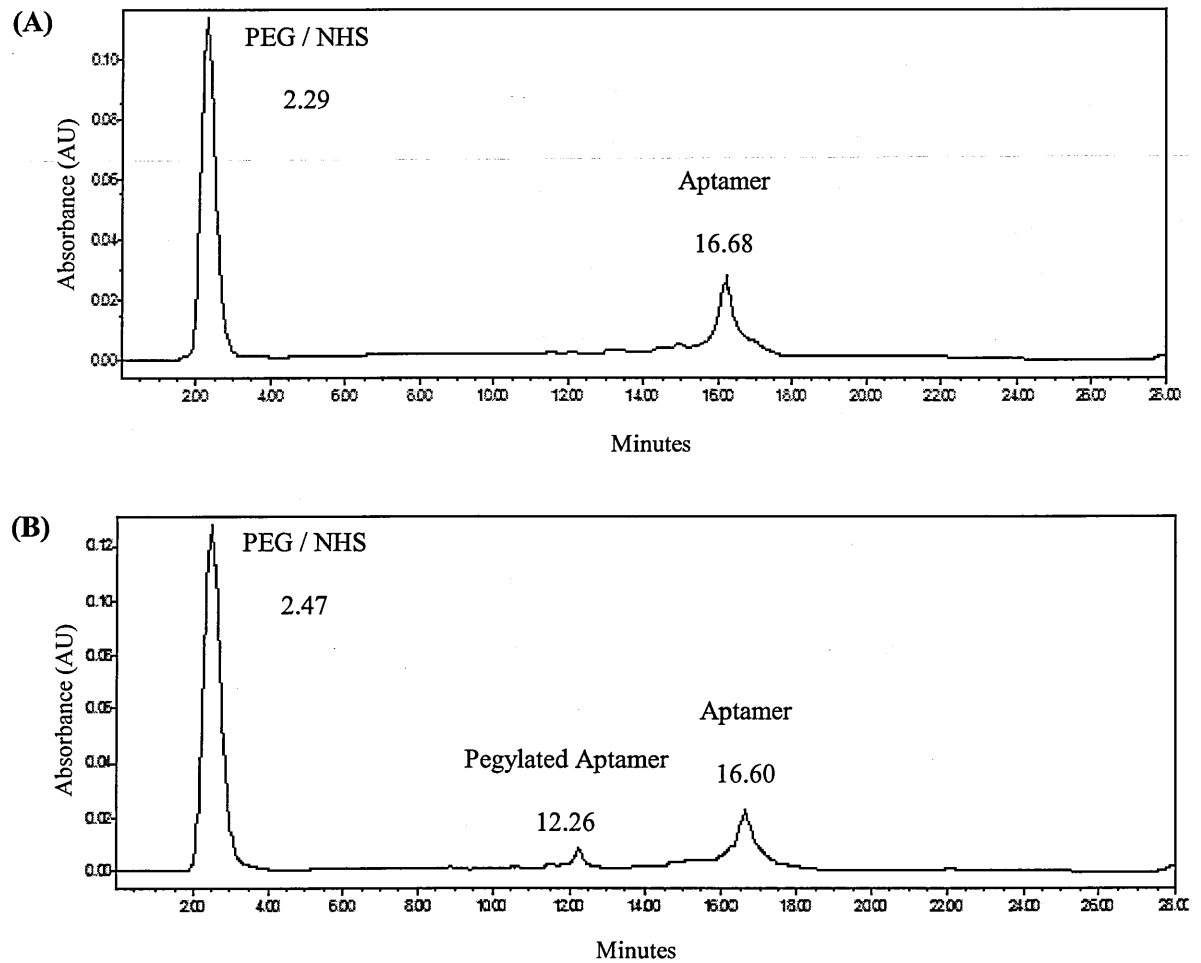


Figure 4.2. HPLC chromatogram of the pegylation reaction mixtures. **(A)** Control reaction: The hydrolysed PEG and NHS elutes at 2.29 minutes. The unreacted aptamer elutes at 16.68 minutes. **(B)** Pegylation reaction: The peak at 2.47 minutes is PEG and NHS. The aptamer elutes at 16.6 minutes and the pegylated aptamer has an R_t of 12.26 minutes.

The formation of the product with a R_t of 12.26 minutes was increased when the pegylation reaction was performed by reacting amino modified aptamer with 40 molar excess of the NHS activated PEG in a 50: 50 (v/v) mixture of 53.4 mM sodium carbonate buffer (pH 10) and DMF at 37°C for 1 hour (figure. 4.3). The pegylated aptamer elutes at 12.09 minutes with a yield of 36 %.

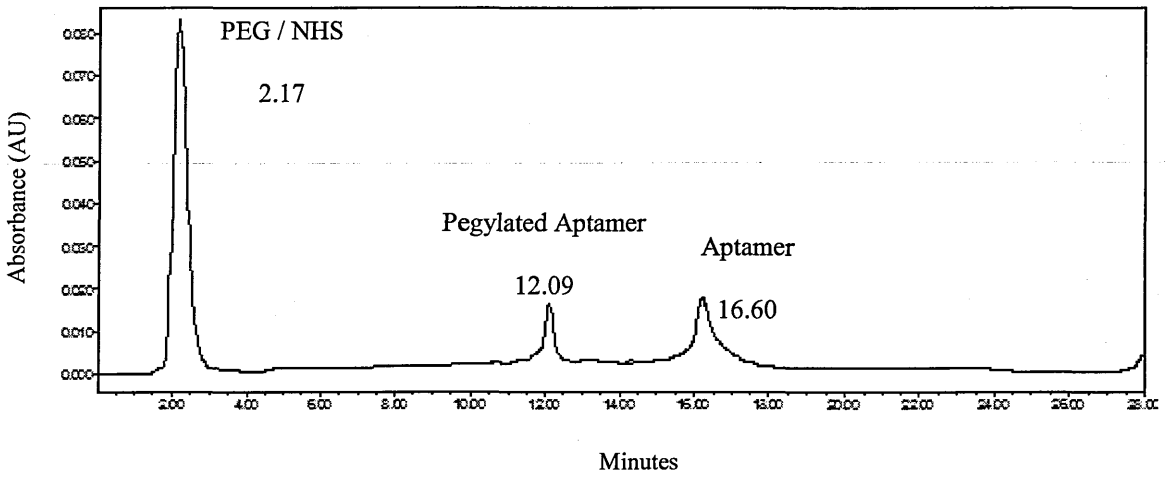


Figure 4.3. HPLC chromatogram of the pegylation reaction: addition of DMF. The hydrolysed PEG and NHS elutes at 2.17 minutes. The unreacted aptamer elutes at 16.60 minutes and the pegylated aptamer elutes at 12.09 minutes.

The increase in temperature from 37°C to 60°C and the percentage of DMF volume to 85 % and reducing the buffer volume to 15 % also led to a small increase in the product (figure. 4.4). The conjugated aptamer elutes at 12.04 minutes with a yield of 42 % of the pegylated aptamer.

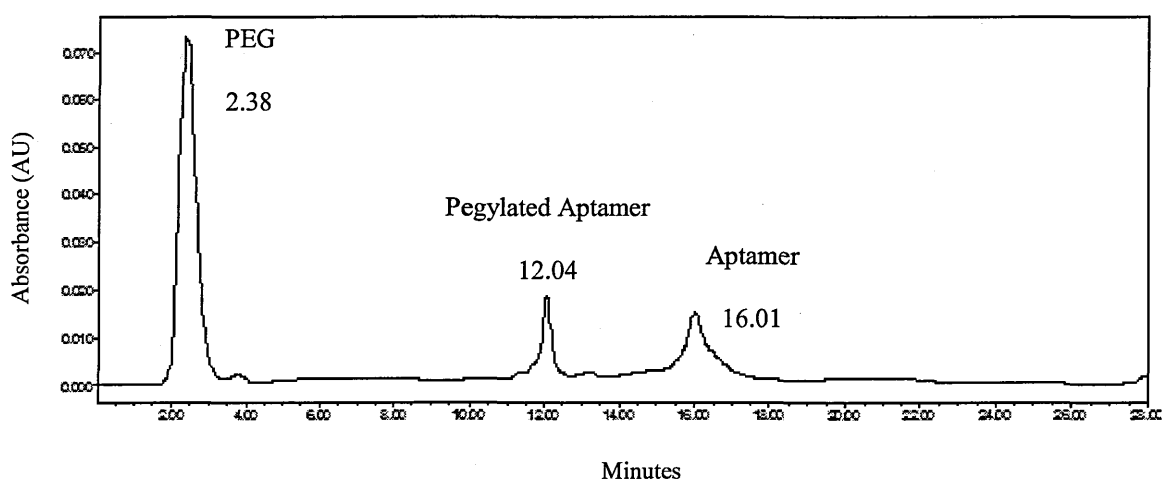


Figure 4.4. HPLC chromatogram of the pegylation reaction: 85 % DMF at 60°C. The hydrolysed PEG and NHS elutes at 2.38 minutes. The unreacted aptamer elutes at 16.01 minutes and the pegylated aptamer elutes at 12.04 minutes.

As the increase of DMF and temperature did not make a significant difference, the previous reaction was repeated with the addition of the 40 molar excess NHS-activated PEG to the aptamer solution in portions every 20 minutes (figure. 4.5) whereby the conjugated aptamer elutes at 12.26 minutes with a yield of 79 %.

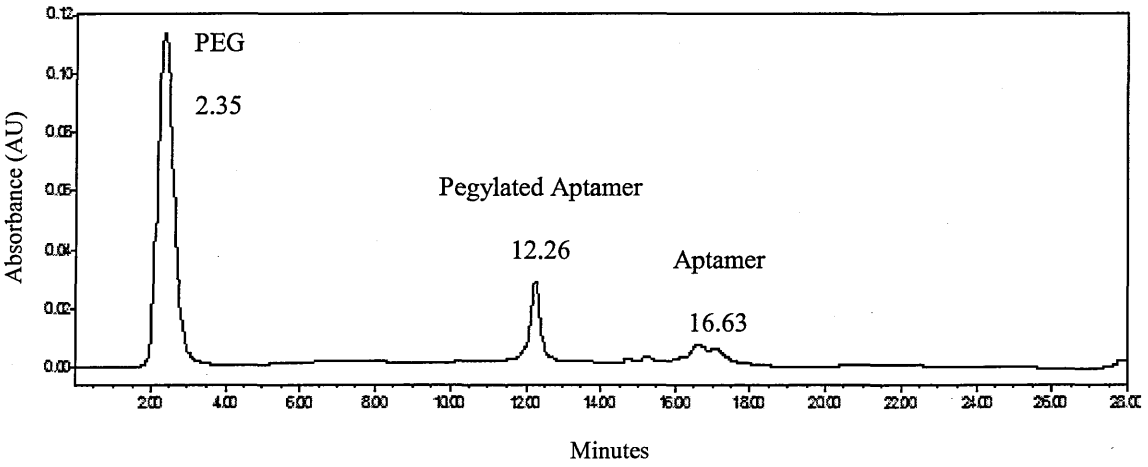


Figure 4.5. HPLC chromatogram of the pegylation reaction: NHS-PEG added in portions. The hydrolysed PEG and NHS elutes at 2.35 minutes. The unreacted aptamer elutes at 16.63 minutes and the pegylated aptamer elutes at 12.26 minutes.

This pegylation condition resulting in a 79 % yield of the conjugated aptamer was employed for the pegylation of all six truncated MUC1/Y aptamer (S11b, S11c, S51b, S51c, S75b and S75c). The pegylated aptamers were purified and isolated by anion exchange HPLC. The isolated pegylated aptamers were analysed on a 2 % agarose gel (figure 4.6). In gel A, the pegylated aptamers S11b and S11c in lanes 3 and 4 respectively have migrated close to the loading wells whilst the free aptamer has migrated in line with the 25 bp mark. In gel B, lane 2, the free PEG is observed migrating towards the cathode. This analysis confirms that the high molecular weight molecule retarded in the gel does contain the aptamer and the PEG, observed as a band as a result of EtBr and BaClI staining.

The isolated aptamer was also analysed by MALDI mass spectrometry for a concluding confirmation (figure 4.7). The free PEG is observed due to the fragmentation of the conjugated molecule and is indicated by its molecular weight with a peak at 33204 Da. The conjugated aptamer is denoted by the peak 40484 Da.

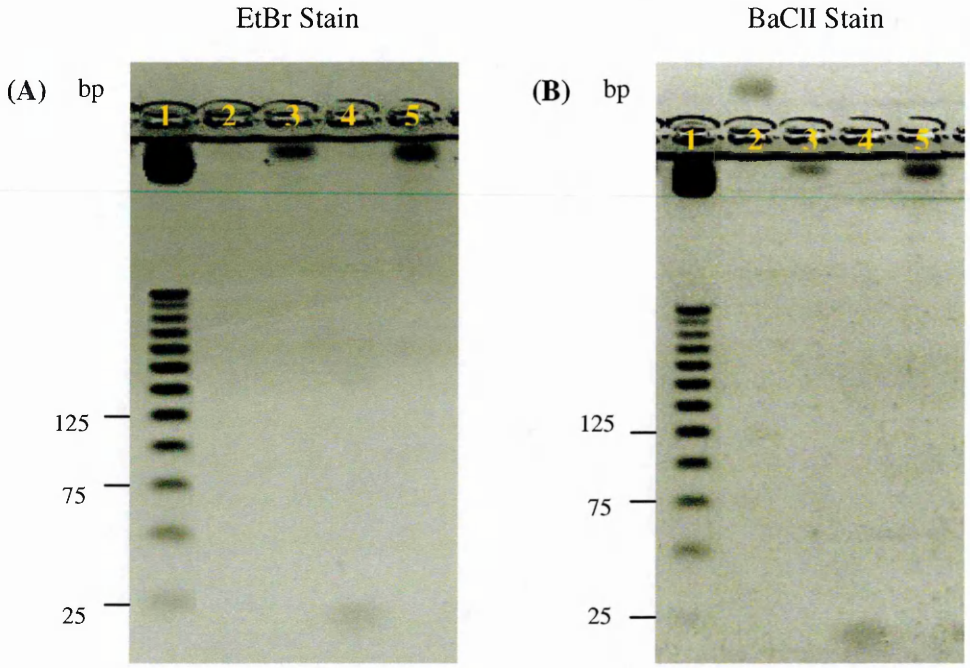


Figure 4.6. Analysis of the isolated conjugated aptamer. 2% agarose gel: (**1**) 25 bp ladder, (**2**) free PEG, (**3**) pegylated S11b, (**4**) free aptamer, (**5**) pegylated S11c.

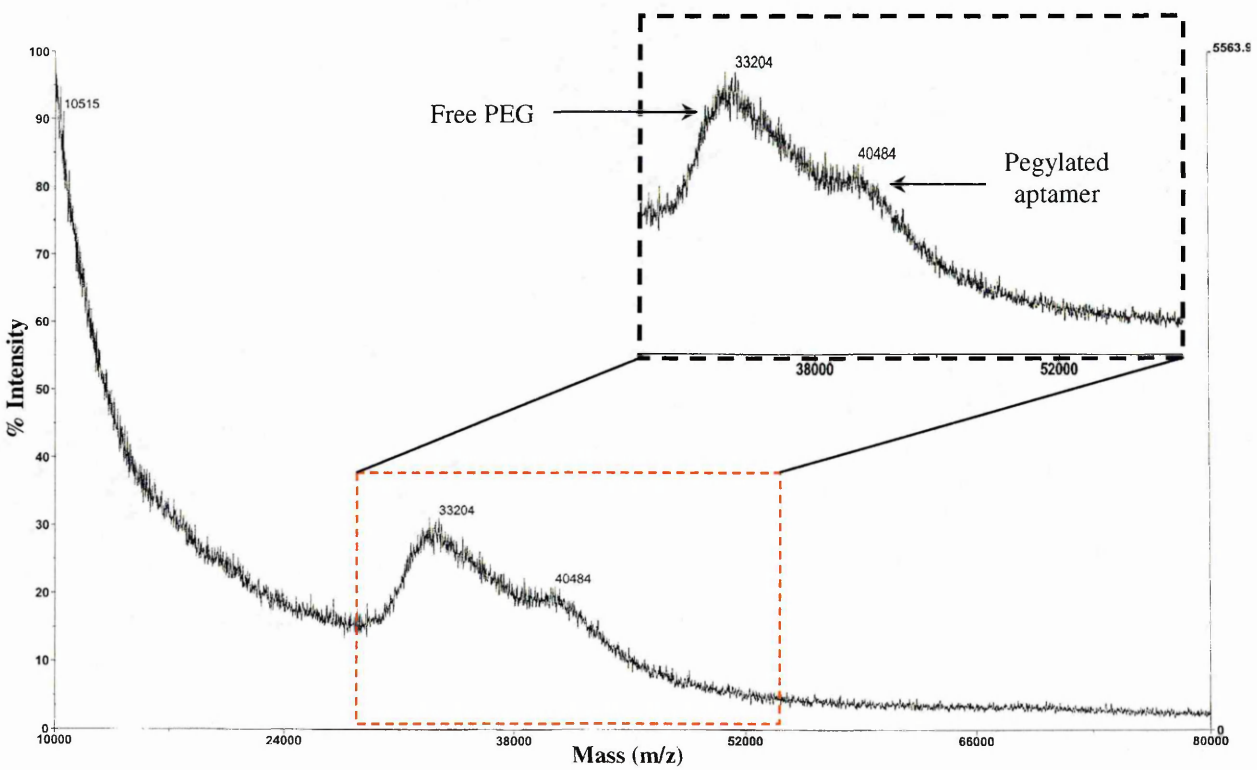


Figure 4.7. MALDI mass spectrum of the pegylated aptamer. Pegylated aptamer: 40484 Da, Free PEG: 33204 Da.

For the purpose of fluorescence *in vitro* cell binding studies, the six truncated aptamers were labelled on the 5′ terminus with a fluorescent Cy3 dye, via phosphoramidite chemistry, and conjugated to the 32 kDa PEG on the 3′ end using the optimised conditions to yield *ca.* 80 % of pegylated aptamer. The Cy3 labelled pegylated aptamers were purified and isolated by anion exchange HPLC and analysed on a 2% agarose gel (figure 4.8) where in lane 1, the conjugated aptamer has migrated close to the well indicated by a bright pink band. The free PEG is observed as rusty coloured band in lane 2 migrating towards the cathode.

Figure 4.8. Analysis of isolated Cy3 labelled conjugated aptamer. 2% agarose gel: (1) pegylated aptamer **(2)** free PEG.

A total of 12 truncated MUC1/Y aptamers were pegylated on the 3′ end of which a set of six aptamers were non-labelled and the other set of 6 aptamers were labelled with the Cy3 fluorescent dye on the 5′ terminus.

4.3 CONCLUSION

The short versions of the MUC1/Y oligonucleotid with a molecular weight of *ca.* 8 kDa have been conjugated to a large molecular weight polymer, PEG of 32 kDa with the objective to increase the molecular weight of the aptamers above the minimum glomerular filtration threshold. Increasing the overall molecular weight of the aptamers has shown to appreciably reduce the rate of renal clearance and therefore extend their *in vivo* circulatory half-life. The most described and widely used method in literature for aptamer pegylation has involved an amino coupling with the 3′ amino modified aptamers and the activated NHS-ester moiety of PEG, which was also employed for the pegylation of the MUC1/Y

aptamers. The initial conditions applied for aptamer conjugation was based on those presented in literature however, with these conditions aptamer pegylation was not achieved. Therefore, following a series of modifications involving changing parameters such as buffer pH, addition of solvent, temperature and PEG ratios, have led to the optimised pegylation conditions with a yield of *ca.* 80% of pegylated aptamer. Optimisation of the pegylation conditions increased the aptamer conjugation yield from 13 % to 79 %. This 66 % improvement in aptamer pegylation yield was clearly observed in the HPLC chromatograms as the additional peak at *ca.* 12.05 was increasing with each modification. The addition of DMF alone has led to a 23 % increase in the formation of the conjugated aptamer, whilst increasing the volume of DMF and temperature to 60°C, a rise of only 6% in the yield was observed. The greatest increase in the formation of the aptamer conjugate was 37 % as a result of adding PEG to the reaction mixture in portions. Thus, most significant improvements are largely attributed to the addition of the PEG in portions and the incorporation of the DMF. Both of these alterations contribute towards a single factor of reducing the rate of hydrolysis and increasing the portion of activated PEG available to the amino modified aptamer for coupling to take place. The final conditions used to pegylate the twelve aptamers (unlabelled and Cy3 labelled) involved reacting a 3′ amino modified aptamers with 43 molar excess of the NHS activated PEG, which was added in portions, in a 15 : 85 (v/v) mixture of 53.4 mM sodium carbonate buffer (pH 10) and DMF at 60°C over a period of two hours. The formation of the pegylated aptamer has been confirmed by HPLC, gel electrophoresis (using EtBr and BaClI staining) and MALDI mass spectrometry.

Analysis of aptamer pegylation via BaClI staining on a 2 % agarose gel has shown to be an ineffective method for the detection of aptamer-PEG complex at low concentrations. However, this was not the case when using the BaClI stain on a 7 % native page. Furthermore, gel electrophoresis analysis can present discrepancies in the analysis of the reaction mixture such as the potential conversion of the free aptamer to pegylated aptamer, which are eliminated when analysed using HPLC.

As PEG is a large bulky molecule, the attachment of this polymer to the aptamers may have the potential to affect the specificity and affinity of the aptamers to their target peptides, thus it is essential that the

pegylated aptamers are characterised alongside the unmodified aptamers. The characterisation of the aptamers will determine any potential change in the binding of the aptamers to the biomarker subsequent to conjugation. Therefore, only after establishing whether the pegylated aptamers retain their affinity and specificity for their target can they be taken further for in vivo studies. The fluorescently labelled Cy3 aptamers were pegylated for comparing their potential with the unpegylated aptamers in the *in vitro* cell binding and internalisation studies.

CHAPTER FIVE

QUALITATIVE MUC1/Y APTAMER CHARACTERISATION

5.0 INTRODUCTION

The generation of three particular MUC1/Y aptamers (S11, S51 and S75) from two different selection methods could be an indication of a significant relationship with respect to their affinity and specificity for their MUC1/Y target peptides (10mer and 20mer). However, qualitative and quantitative characterisation studies are necessary to reveal the affinity and specificity of these aptamers. The determination of aptamer affinity will provide information on the strength of the interaction between the aptamer and its target, which can also be indicated by the stability of the aptamer-target complex. Whilst, specificity of an aptamer can be determined by evaluating the interaction of aptamers, those selected against the target and those randomly chosen as a control, with the target and other random peptides. To ascertain the affinity and specificity of aptamers towards their target, numerous approaches can be employed, such as nitrocellulose partitioned aptamer binding [65,66], enzyme-linked immunosorbent assay (ELISA) [197,233], surface plasmon resonance (SPR) [234,235], thermal denaturation [66,234], electrophoretic mobility shift assay (EMSA) [233,236], or affinity chromatography [197]. Many biophysical assays require modification of the aptamers. In the case of nitrocellulose partitioned aptamer binding and ELISAs, the aptamer is modified with a radioactive label or biotin respectively, for establishing binding constants. However, such modifications hold the potential to alter the original affinity of the aptamer, therefore binding constants determined from these assays cannot be verified as the true affinity of the unmodified aptamer. To assess the binding affinity and specificity of MUC1/Y aptamers of series a (S11a, S51a and S75a), series b (S11b, S51b and S75b) and series c (S11c, S51c and S75c), the primary focus was placed on employing qualitative biophysical assays such as EMSA, affinity chromatography and thermal denaturation studies which do not require the aptamers under investigation to be modified.

5.1 PRINCIPLE OF THE EMSA

The electrophoretic mobility shift assay (EMSA), also known on occasions as the gel or band shift or gel retardation assay, is used for studying complex formation between typically an oligonucleotide and protein [236,237]. Non-denaturing polyacrylamide gel electrophoresis is used to monitor the separation of the complex from the unbound components. The migration of the molecules through the gel is mainly

dependent upon their size and charge, although the conformation of a molecule can also have an influence on the mobility. A protein-oligonucleotide complex will have a lower mobility than the unbound oligonucleotide and protein due to its overall molecular weight. However, in the absence of any complex formation, retardation of the oligonucleotide will not be observed and will possess the same mobility as the free oligonucleotide used as a control. The critical element in this technique for optimal results is the pore size of the gel which is controlled by the concentration of the acrylamide and bisacrylamide used to form the gel. The pore size of the gel is tailored to accommodate not only the molecular weight of the unbound protein and oligonucleotide but also the complex. A small pore size of the gel will prevent the large protein-oligonucleotide complex from entering the gel, whilst a very large poor size may not participate in the retardation of the complex thus, resulting in an insufficient separation of the complex from the unbound molecules. This EMSA technique has been applied to resolve the aptamer-peptide complexes from the unbound aptamer.

5.2 EMSA WITH THE MUC1/Y APTAMERS

Each aptamer (S11a, S11b, S11c, S51a, S51b, S51c, S75a, S75b, and S75c) was incubated with the peptides (10 mer and 20 mer) in a 1: 1000 molar ratio in binding solution. The incubation of the aptamer-peptide solutions was carried out using two different conditions: 1) incubating the solution for 1 hour at ambient temperature with gentle shaking, 2) heating the solution at 95°C for 5 minutes, followed by incubating the solution for 1 hour at ambient temperature with gentle shaking.

Subsequent to incubation, the samples were analysed on native PAGE and the gels were post stained with ethidium bromide. Figure 5.0 shows the analysis of the solution mixture of aptamers S11a, S51a and S75a (72 nt long) on a 14 % native PAGE. The free aptamer, used as a control is seen are lanes 5 and 8. The S11a- 10 mer, S11a- 20 mer, S51a- 10 mer and S51a- 20 mer mixtures are seen are lanes 4, 6, 7 and 9 respectively. The band in lane 6 (S11a- 20 mer mixture) appears to be migrating just very slightly faster than the free aptamer and band 4 (S11a- 10 mer mixture). The band in lane 7 appears to be migrating vaguely faster than the free aptamer (lane 8). The samples which were not exposed to heat show the

presence of the aptamers at *ca.* 75 bp with very low band shift patterns due to the low molecular weight of the 10 mer and 20 mer peptides of *ca.* 1 kDa and 2 kDa. When the aptamer-peptide mixtures heated at 95°C were analysed by PAGE no bands were detected, lanes 1, 2, 10 and 11 (Figure 5.0), suggesting that after the thermal denaturation, the aptamer is blocked from re-forming a secondary structure, necessary for the detection by the EtBr dye which is unable to intercalate with the aptamers to result in fluorescence during the exposure to UV light. Similar shift patterns were also observed for the S75a aptamer (appendix 2.1).

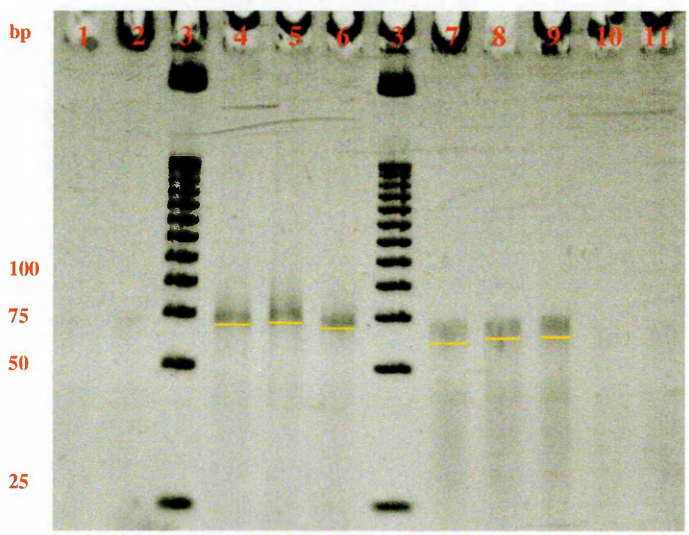


Figure 5.0. Analysis of the S11a and S51a-peptide mixture. 14% native PAGE: S11a: (1) S11a + 10 mer at 95°C, **(2)** S11a + 20 mer at 95°C, **(3)** 25 bp ladder, **(4)** S11a + 10 mer, **(5)** free S11a, **(6)** S11a + 20 mer, **S51a: (7)** S51a + 10 mer, **(8)** free S51a, **(9)** S51a + 20 mer, **(10)** S51a + 10 mer at 95°C, **(11)** S51a + 20 mer at 95°C. The peptides do not stain with ethidium bromide and therefore were not analysed with the samples.

The solutions of 25 nt long MUC1/Y aptamer (S11b S11c, S51b, S51c, S75b and S75c) and peptides were analysed on a 18 % native PAGE (figure 5.1). The free aptamer, used as a control is seen in lanes 5 and 8. The S11b- 10 mer, S11b- 20 mer, S11c- 10 mer and S11c- 20 mer mixtures are seen in lanes 4, 6, 7 and 9 respectively. When the heated at 95°C aptamer-peptide mixtures were analysed, lanes 1, 2, 10 and 11, no bands were detected. For both aptamers (S11b and S11c) distinguishable shifts were not obtained whilst no aptamer bands are observed with the heated aptamer peptide solutions. Similar shift patterns were also observed for the aptamers S51b, S51c, S75b and S75c (appendix 2.2 - 2.4).

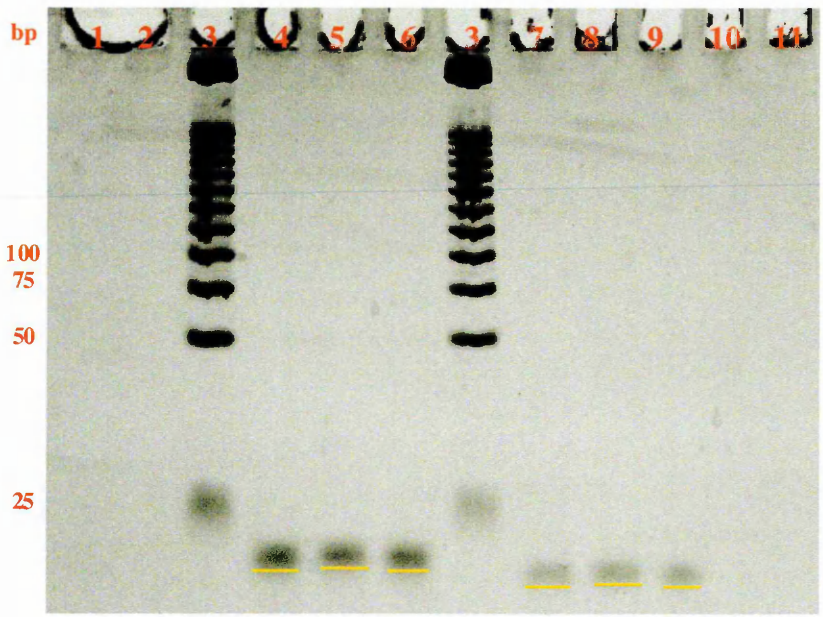


Figure 5.1. Analysis of the S11b and S11c-peptide mixture. 18% native PAGE: S11b: (1) S11b + 10 mer at 95°C, **(2)** S11b + 20 mer at 95°C, **(3)** 25 bp ladder, **(4)** S11b + 10 mer, **(5)** free S11b, **(6)** S11b + 20 mer, **S11c: (7)** S11c + 10 mer, **(8)** free S11c, **(9)** S11c + 20 mer, **(10)** S11c + 10 mer at 95°C, **(11)** S11c + 20 mer at 95°C. The peptides do not stain with ethidium bromide and therefore were not analysed with the samples.

EMSA's of the control aptamers with the 10 mer and 20 mer peptides at ambient temperature and heating to 95°C, were performed using aptamers S11b-smd, S11c-smd and S75b-smd (scrambled versions of the S11b, S11c and S75b respectively), SP68 (selected against a non-MUC1 tumour target peptide) and S2.2 (selected against MUC1)[197]. The sequences and the predicted secondary structures of the control aptamers are in appendix 2.0. The mixtures were analysed on an 18 % native PAGE (figure 5.2). The free aptamers correspond to the bands in lanes 5 and 8. The bands of S11b-smd-10 mer, S11b-smd-20 mer, S11c-smd-10 mer and S11c-smd-20 mer mixtures were present in lanes 4, 6, 7 and 9 respectively. Both of the scrambled aptamers (S11b-smd and S11c-smd) appear to have a faster mobility than the 25 bp marker and compared to the MUC1/Y aptamers S11b and S11c. Also in both cases the aptamer-20mer band is migrating slightly faster than the free aptamer and the aptamer-10mer band. The aptamer-peptide mixtures heated at 95°C were analysed in lanes 1, 2, 10 and 11, where bands around the 25 bp mark were not observed. EMSA's of S75b-smd, SP68 and MUC1-S2.2 control aptamers also show either no or insignificant band shifts (appendix 2.3 and 2.5).

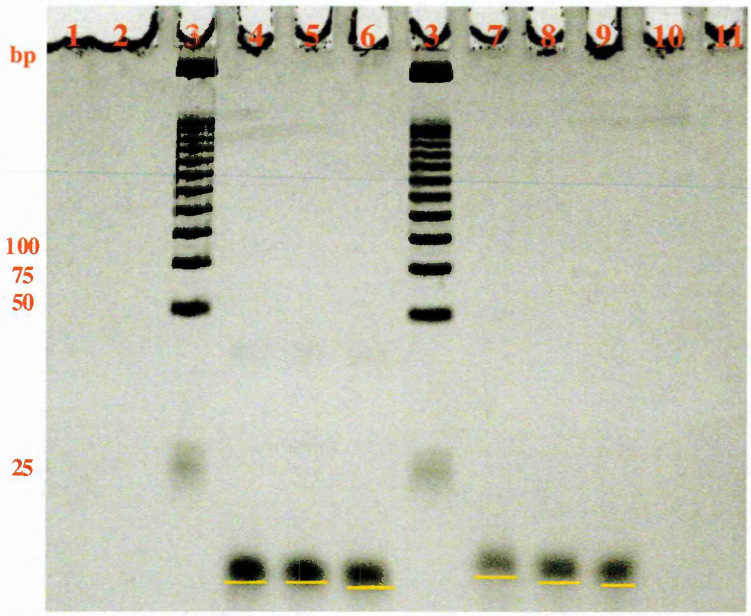


Figure 5.2. Analysis of the S11b-smd and S11c-smd-peptide mixture. 18% native PAGE: S11b-smd: (**1**) S11b-smd + 10 mer at 95°C, (**2**) S11b-smd + 20 mer at 95°C, (**3**) 25 bp ladder, (**4**) S11b-smd + 10 mer, (**5**) free S11b-smd, (**6**) S11b-smd + 20 mer, **S11c-smd:** (**7**) S11c-smd + 10 mer, (**8**) free S11c-smd, (**9**) S11c-smd + 20 mer, (**10**) S11c-smd + 10 mer at 95°C, (**11**) S11c-smd + 20 mer at 95°C. Smd = scrambled aptamer. The peptides do not stain with ethidium bromide and therefore were not analysed with the samples.

The pegylated aptamers S11b-PEG, S11c-PEG, S51b-PEG, S51c-PEG, S75b-PEG and S75c-PEG were also incubated with the two peptides under the previously described conditions and the mixtures were analysed on a 7 % native PAGE (figure 5.3). All the pegylated aptamers were significantly retarded in the gel as expected and have migrated slightly further away from the loading wells. The free aptamer, used as a control is seen in lanes 5 and 8. The S11b-PEG-10 mer, S11b-PEG-20 mer, S11c-PEG-10 mer and S11c-PEG-20 mer mixtures are seen in lanes 4, 6, 7 and 9 respectively. For both aptamers, the 20 mer band was migrating slightly faster than the 10 mer band and the free aptamer. The pegylated aptamer-peptide mixtures heated at 95°C were analysed in lanes 1, 2, 10 and 11 and no bands were detected when stained with EtBr, however the presence of the pegylated aptamer was indicated by the rusty coloured bands when stained with BaCll solution. Comparable shift patterns were also observed for the pegylated aptamers S51b-PEG, S51c-PEG, S75b-PEG and S75c-PEG (appendix 2.6 - 2.7).

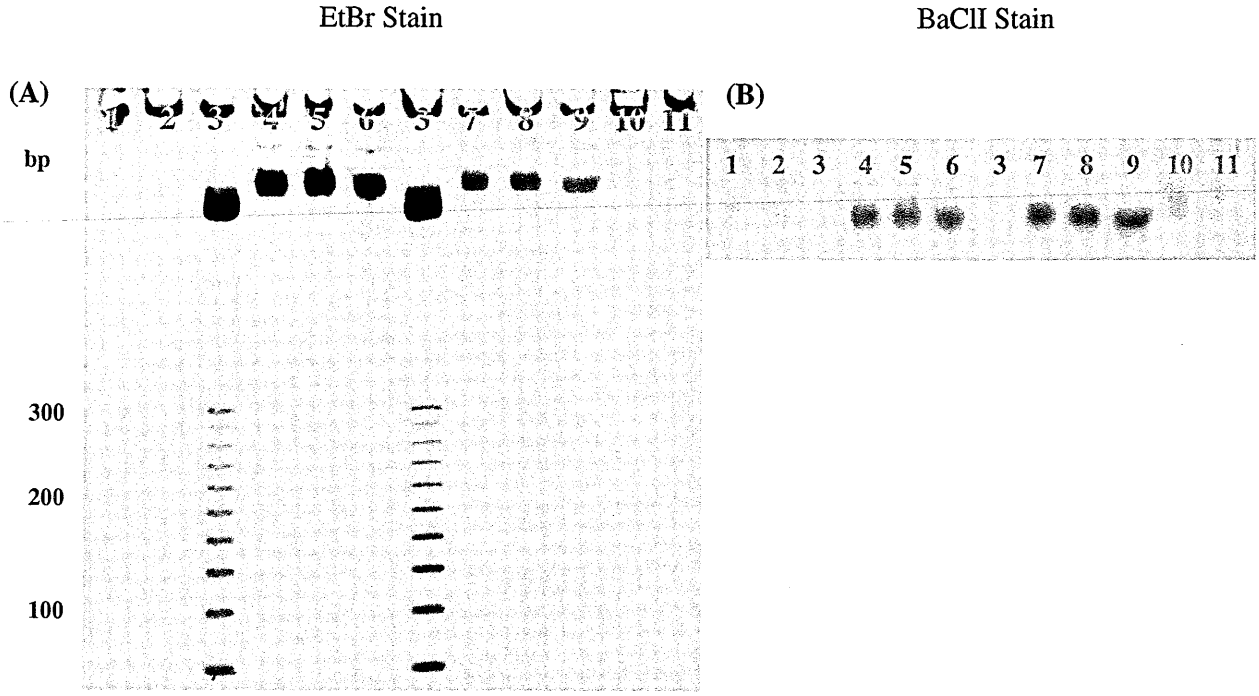


Figure 5.3. Analysis of the pegylated S11b and S11c-peptide mixture. 7 % native PAGE: (A) S11b-PEG: (**1**) S11b-PEG + 10 mer at 95°C, (**2**) S11b-PEG + 20 mer at 95°C, (**3**) 25 bp ladder, (**4**) S11b-PEG + 10 mer, (**5**) free S11b, (**6**) S11b-PEG + 20 mer, **S11c-PEG:** (**7**) S11c-PEG + 10 mer, (**8**) free S11c, (**9**) S11c-PEG + 20 mer, (**10**) S11-PEG c + 10 mer at 95°C, (**11**) S11c-PEG + 20 mer at 95°C. **(B)** Same as gel A, but stained with BaCll for staining of PEG. The peptides do not stain with ethidium bromide and therefore were not analysed with the samples.

All analysed MUC1/Y aptamers (full length, truncated, and pegylated) and the control aptamers exhibit negligible band shifts when incubated with the 10 mer and 20 mer peptides since the MUC1/Y peptides have a low molecular weight (1 – 2 kDa) compared to the full length (22 kDa), truncated (8 kDa) and pegylated (40 kDa) aptamers. The EMSA is a not a suitable technique to detect the aptamer-peptide complex formation due to the inability to resolve the very small molecular weight differences. Moreover, all the aptamers including the control aptamers, when heated to 95°C for 5 minutes appeared to interact with the peptides which interfered with the formation of the secondary structure of the aptamer. This effect was indicated in the gel by the disappearance of the aptamer bands after staining with EtBr, an intercalating dye. This suggests that non-specific aptamer binding to the target when fully denatured (presence of no secondary structure) is likely to take place.

5.3 PRINCIPLE OF THE AFFINITY CHROMATOGRAPHY TECHNIQUE

Affinity column chromatography is mostly employed for aptamer selection against a target but can also be used for the assessment of aptamer specificity and affinity for its target. The ionic strength required to disrupt the target-oligonucleotide complex is correlated to the affinity and specificity of the oligonucleotides for their targets. In this system, the target is immobilised onto a column containing an NHS activated sepharose matrix to which the oligonucleotide solution is then loaded and incubated with the stationary target. The elution is performed using a solution gradient of increasing salt concentrations. Any excess or unbound oligonucleotide is eluted from the column at the lowest concentration of salt solution. The remaining bound aptamer is eluted by the mobile phase of increasing polarity. Each eluted fraction corresponding to the different molar salt concentrations is analysed by UV spectroscopy to detect the presence of oligonucleotide. The salt concentrations required to elute each aptamer from the target are compared to determine the affinity and specificity of the aptamers for their target as described in the literature [197].

5.4 AFFINITY CHROMATOGRAPHY WITH THE MUC1/Y APTAMERS

The MUC1/Y peptides (10 mer and 20 mer) were individually immobilised onto a NHS activated sepharose matrix of a 1 mL HiTrap affinity chromatography column. Each aptamer, S11 series (S11a, S11b, S11b-smd, S11b-PEG, S11c, S11c-smd, and S11c-PEG), S51 series (S51a, S51b, S51b-PEG, S51c, S51c-PEG and SP68), and S75 series (S75a, S75b, S75b-smd, S75b-PEG, S75c and S75c-PEG) in binding solution, was added to the affinity column and incubated at ambient temperature for 1 hour. The column was then eluted with elution solutions (1 mL each) of increasing NaCl concentrations ranging from 0.2 M – 1.6 M at 0.2 M increments. Subsequently, to regenerate the affinity chromatography columns, these were washed with a 3 M sodium thiocyanate solution and re-equilibrated with binding solution. Each eluted fraction was analysed by UV-Vis spectroscopy measuring the absorbance of the aptamer at 260 nm. The UV absorbance of each eluted aptamer was subtracted from the initial absorbance of the aptamer prior to loading on to the column to ascertain the presence of the aptamer remaining on the affinity column. As indicated in the table shown in appendix 2.8, the UV measurements

for all the aptamers analysed showed the majority of the aptamer had eluted at the lowest NaCl concentration of 0.2 M and the bound aptamers could not be eluted even at the highest NaCl concentration of 1.6 M. NaCl. Concentrations above 1.6 M were not used to avoid the risk of dehydrating the matrix of the chromatography column. Figure 5.4 shows the analysis of the S11 aptamer series. A significant difference of aptamer binding to the two peptides was not observed, suggesting that the length of the target does not appear to have an effect on the binding potential of the aptamers. The binding of the full length aptamer S11a (25.9 % and 29.3 %) is comparable to the truncated S11b (31.6 % and 21.5 %) and the control aptamer S11b-smd (27.1 % and 20 %) to the 10 mer and 20 mer peptide respectively. However, the binding of the truncated S11c (39.1 % and 35.4 %) is moderately higher than S11a, S11b and the control aptamer S11c-smd (28.1 % and 25.0 %) to the 10 mer and 20 mer peptide respectively. Interestingly, the S11b-PEG (47.4 % -20 mer) and S11c-PEG (77.7 % and 83.2 %) has shown substantial increase in binding for the 10 mer and 20 mer peptides respectively compared to that of the unmodified aptamers.

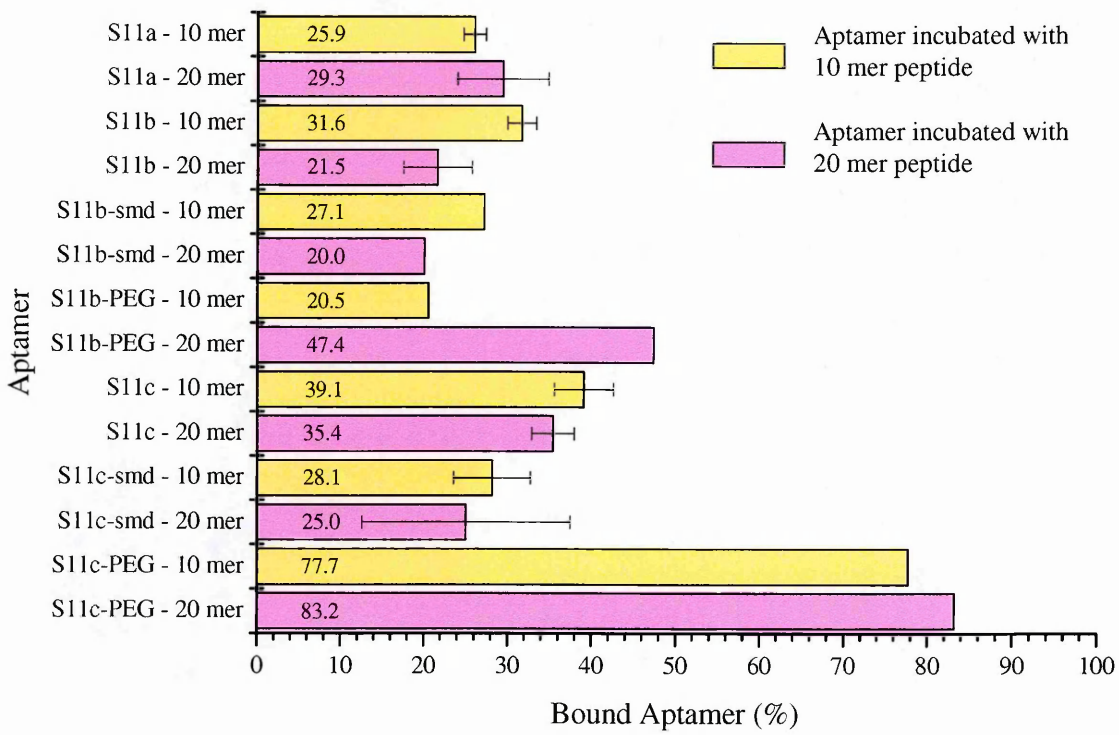


Figure 5.4. Aptamer binding of S11 aptamer series. Each bar represents the percentage of bound aptamer subsequent to the 1.6 M NaCl elution of the affinity column for each MUC1/Y peptide (10 mer and 20mer). Values are expressed as mean % (n = 3), with the exception of the pegylated aptamers and S11b-smd (n = 1).

Figure 5.5 shows the analysis of the S51 aptamer series and the SP68 control aptamer with the MUC1/Y peptides. In this series, each aptamer has a greater percentage of binding to the 10 mer peptide than the 20mer peptide and the truncated aptamers (S51b and S51c) bind less than the full length aptamers (S51a). The binding of the full length aptamer S51a (50 % and 45.2 %) is notably higher than the truncated S51b (29.5 % and 24.5 %) and truncated S51c (33.8 % and 25.7 %), whilst the binding of the control aptamer, SP68 (a non-MUC1 tumour target peptide) (26.4 % and 25.4 %) did not show substantial difference in binding compared to truncated aptamers. Also, like in the S11 series, S51b-PEG (79.8 % and 59.8 %) and S51c-PEG (89.3 % and 76.2 %) have shown a significant increase in the binding to the 10mer and 20 mer peptides respectively compared to the unmodified aptamer.

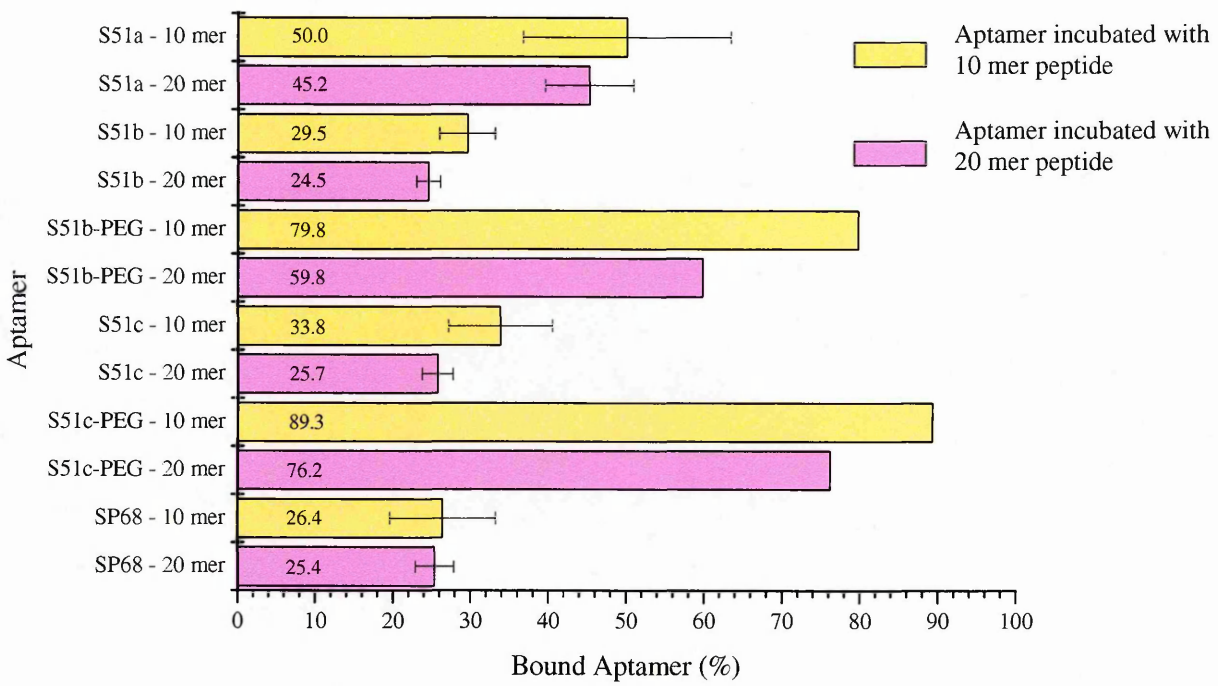


Figure 5.5. Aptamer binding of S51 aptamer series. Each bar represents the percentage of bound aptamer subsequent to the 1.6 M NaCl elution of the affinity column for each MUC1/Y peptide (10 mer and 20mer). Values are expressed as mean % (n =3), with the exception of the pegylated aptamers (n = 1).

Figure 5.6 shows the analysis of the S75 aptamer series. A significant difference in the aptamer binding to the two peptides did not seem apparent as the percentage of aptamer bound to both peptides was relatively similar, as observed in the analysis of the S11 aptamer series. The binding of the full length aptamer S75a (30.5 % and 31.4 %) is comparable with the binding of the truncated aptamers S75b (33.3

% and 24.7 %) and S75c (25.4 % and 28.8 %) to the 10 mer and 20 mer peptides respectively. The binding of the control aptamer S75b-smd (28.9 % and 21.6 %) to the MUC1/Y peptides is slightly reduced by 4.4 % and 3.1 % respectively. Furthermore, like in the pervious aptamer series, the enhanced binding of the pegylated aptamers S75b-PEG (41 % and 37.2 %) and S75c-PEG (54.4 % and 55.2 %) is shown for the 10 mer and 20 mer peptides respectively compared to the unmodified aptamers, though in this case it is perhaps not as pronounced as it has been for the two previous aptamer series.

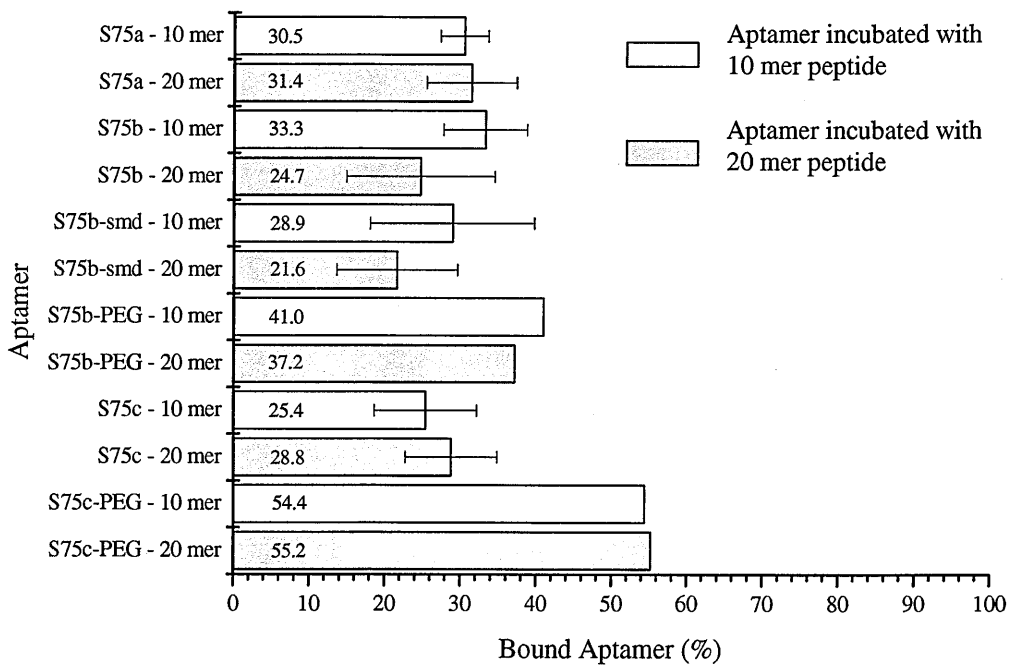


Figure 5.6. Aptamer binding of S75 aptamer series. Each bar represents the percentage of bound aptamer subsequent to the 1.6 M NaCl elution of the affinity column for each MUC1/Y peptide (10 mer and 20mer). Values are expressed as mean % (n = 3), with the exception of the pegylated aptamers (n = 1).

Considerable changes in the percentages of aptamer bound to the peptides were not observed between the full length and the truncated aptamers, or with the control aptamers. The affinity chromatography technique has shown that for all the aptamers analysed, 0.2 M NaCl fraction contained DNA being possibly the excess and unbound aptamer. Aptamer elution in any other salt concentrations was not observed. For aptamer series S11 and S75 it was observed that the length of the peptides did not signify an importance in the binding of the aptamer. Although, this was not the case for the S51 aptamer series, where binding of the aptamers to the 10 mer peptide was enhanced compared to the 20 mer peptide. The

S51a full length aptamer has greater binding to both MUC1/Y peptides compared to the S11a and S75a. Also, evident from the results for all three aptamer series was the remarkable increase in binding of the pegylated aptamers to the peptides compared to the unmodified aptamers and to the full length aptamers. This could be an implication that the large PEG polymer may affect the structure of the aptamer which has resulted in enhanced binding to the peptides. The specificity of the aptamer for the peptides could not be identified as the percentage of control aptamer bound to the peptides was comparable to the aptamers selected against the MUC1/Y peptides.

5.5 PRINCIPLE OF THE THERMAL DNA DENATURATION TECHNIQUE

Thermal denaturation of oligonucleotides, also referred to as melting, is the dissociation of a double stranded oligonucleotide, resulting in two single strands. This technique utilises heat to break the hydrogen bonds between the bases responsible for the formation of a duplex strand. This can also be applied to single stranded oligonucleotides to assess their conformational changes, whereby the breaking of hydrogen bonds of base-pairs, involved in the formation of the stem structure, is monitored. The melting temperature (T_m) is the temperature at which 50 % of the oligonucleotide secondary structure has been denatured and the remaining 50 % still retains its original conformation. Melting temperature of an oligonucleotide is extensively dependent upon the length and sequence composition of the oligonucleotide. Thus, a sequence rich in guanine-cytosine (G-C) pairs will be more stable than a sequence rich in adenosine-thymine (A-T) pair and therefore higher melting temperatures are observed. The oligonucleotide melting can be monitored by 1) UV absorbance or 2) fluorescence using an intercalating or high resolution melting (HRM) dye. When the UV absorbance is analysed, as the oligonucleotide melts, the signal at 260 nm increases as the temperature rises. In the case of fluorescence analysis, an HRM dye, such as SYBR Green, which binds only to the double stranded section of the oligonucleotide structure, is employed. The binding of the SYBR Green dye to the double stranded oligonucleotide shows fluorescence at 520 nm when excited at 497 nm. During thermal denaturation, as the oligonucleotide melts, the dye is dissociated free into solution, resulting in a decrease in the fluorescence emission. In both methods of analysis, UV and fluorescence, the melting temperature provides an indication of the stability of the secondary structure of an oligonucleotide. Thus, the

determination of the melting temperature can be employed to assess the effects of any modifications made to the oligonucleotide, i.e. whether the alterations are stabilising or weakening the bonds which form the secondary structure.

5.6 THERMAL DENATURATION OF THE MUC1/Y APTAMERS

For the determination of the thermal denaturation of the MUC1/Y aptamers, the fluorescence analysis approach was employed using the HRM SYBR green dye, as this technique is considerably more sensitive than the UV absorbance method. Each aptamer of the S11 series (S11a, S11b, S11b-smd, S11b-PEG, S11c, S11c-smd, and S11c-PEG), the S51 series (S51a, S51b, S51b-PEG, S51c, and S51c-PEG), the S75 series (S75a, S75b, S75b-smd, S75b-PEG, S75c and S75c-PEG) and the additional control aptamers SP68 (selected against a non-MUC1 tumour target peptide) and S2.2 [197] (selected against MUC1) was incubated with the peptides (10 mer and 20 mer) in a 1: 1000 molar ratio, in binding solution, for 1 hour, at ambient temperature. Upon incubation, the SYBR green dye was added to the aptamer solution and incubated for 30 minutes on ice. This thermal denaturation cycle (TDC) was performed three times from 3°C to 100°C at a rate of 1°C/ min with a gradual renaturation process between each cycle. The emission was detected at 520 nm when excited at 497 nm. Figure 5.7 shows plots for S11a (A) and S11a with the 20mer peptide (B) at each TDC. In the case of S11a (figure 5.7A), a minimal variation in fluorescence was observed between the three TDC. When S11a was incubated with the 20 mer peptide (figure 5.7B) on the other hand, the fluorescence value at the beginning of the analysis is slightly lower compared to the free aptamer (figure 5.7A). However, dramatic changes in fluorescence were observed during the second and the third TDCs where the values at the beginning of the cycle were *ca.* 0.5 and *ca.* 0.2 respectively. This effect suggests that the peptide is inhibiting the renaturation of S11a, resulting in the aptamer to remain mainly in an open conformation. A similar trend of gradual decrease in fluorescence through each cycle was also observed when the 10 mer peptide was used.

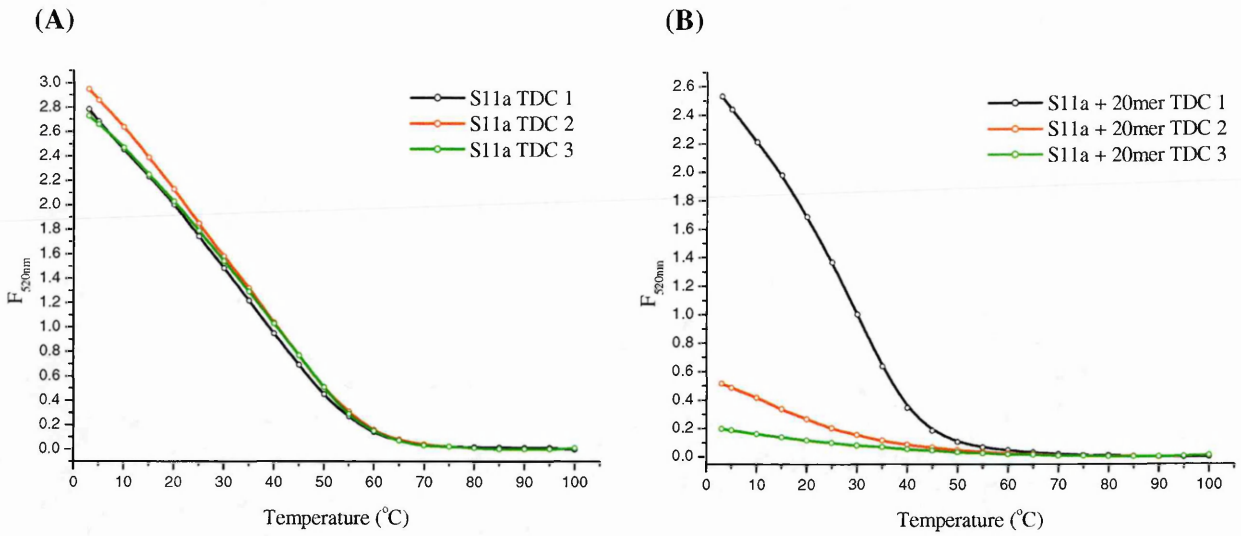


Figure 5.7. Thermal denaturation of S11a. (**A**) Fluorescence melting profiles of S11a alone at 3 TDC. (**B**) Fluorescence melting profiles of S11a with the 20 mer peptide at 3 TDC.

Moreover, all the aptamers, including the control aptamers, exhibited the same behaviour thus indicating that all aptamers appear to bind to the peptide when denatured. Therefore, for the purpose of greater accuracy in the analysis of the thermal denaturation data and for the simplicity of the discussion, only the first TDC for each aptamer with each peptide will be from now analysed and discussed. To determine the effect the peptides have on the conformation of the aptamer, and thus its significance upon binding, the fluorescence melting profiles will be presented as the percentage change in fluorescence (ΔF_{520nm} (%)) *vs* temperature and the melting temperatures (T_m) will be determined as the first derivative. T_m values of each aptamer for all three TDC with both MUC1/Y peptides are presented in the table shown in appendix 2.9. The results from the analysis of the S11 series, without the addition of the peptides are shown in figure 5.8. The truncation of the S11a aptamer to S11b (central variable region) reduces the T_m from 35°C to 21°C, resulting in a loss of thermostability of the aptamer by more than 10°C. The control scrambled version of S11b (S11b-smd) has the same T_m as S11b (21°C), whilst the S11b-PEG appears to possibly induce another stem structure resulting in a second T_m at 40°C. The other possibility could be that the solution may contain a mixture of S11b-PEG with two different structural conformations, one of which melts at 21°C and the other at 40°C. Figure 5.8B shows that truncation of the S11a aptamer to S11c (stem-loop region) has led to an aptamer with two T_m (20°C and 45°C) where the second T_m increases the thermostability of the aptamer by 10°C from 35°C to 45°C. S11c (T_m 20°C and 45°C) also

appears to be more stable than the S11b (21°C). The control scrambled version of S11c (S11c-smd) reverts back to a single T_m of 47°C appearing to be overall more stable than S11b, whilst the pegylated S11c (S11c-PEG) further increases the second T_m by 5°C from 45°C to 50°C. Like the unmodified aptamer S11b and S11c, the S11c-PEG (T_m 20°C and 50°C) appears to be more stable than the S11b-PEG (T_m 20°C and 40°C).

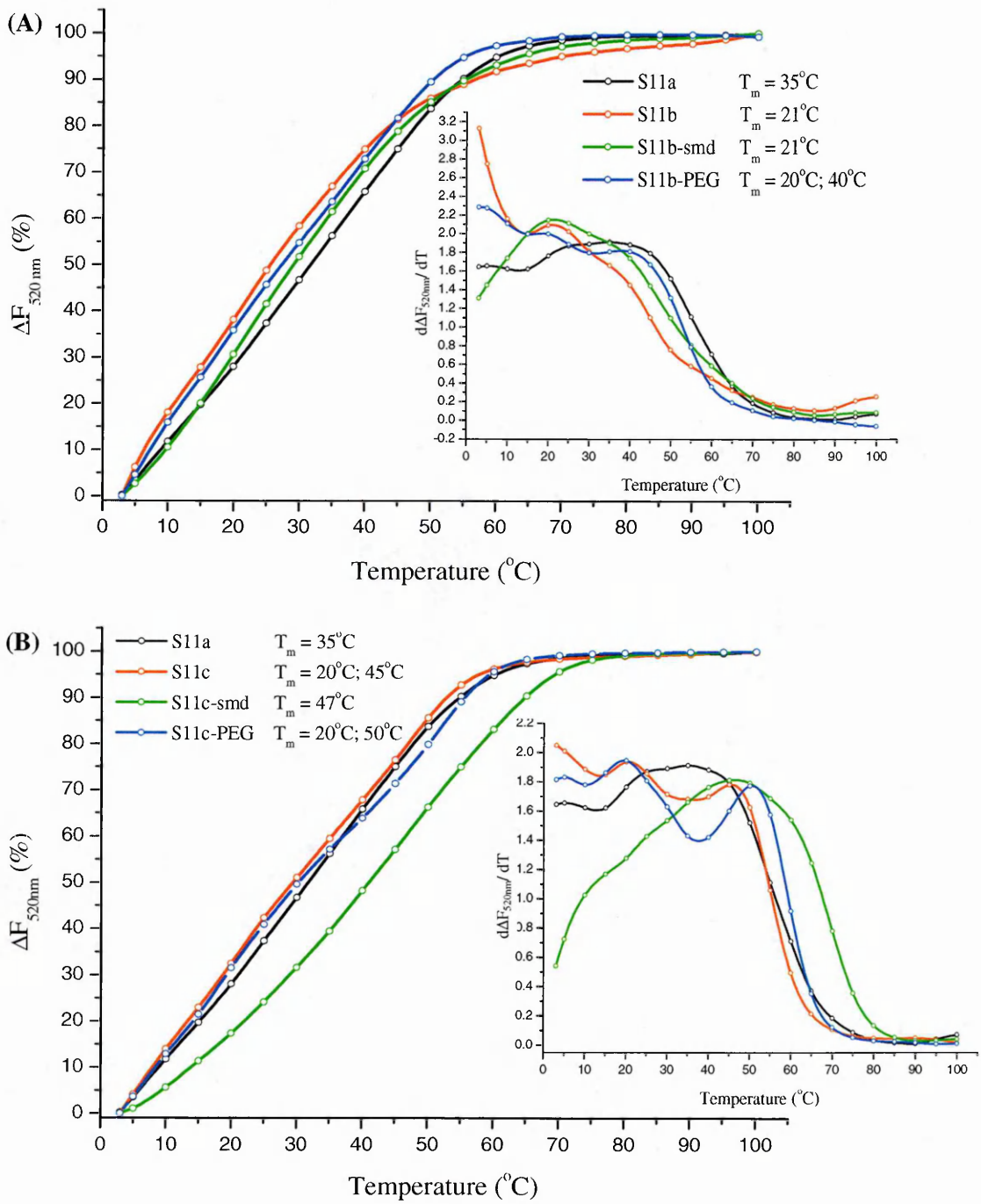


Figure 5.8. Thermal denaturation of S11 series. (**A**) Fluorescence melting profiles of S11a with the S11b versions (**B**) Fluorescence melting profiles of S11a with the S11c versions.

The effect of the MUC1/Y peptides added to the S11a aptamer is shown in figure 5.9. The addition of the 10 mer peptide increases the T_m of the aptamer from 35°C to 43°C, whilst the addition of the 20mer peptide reduces the T_m by 5°C down to 30°C.

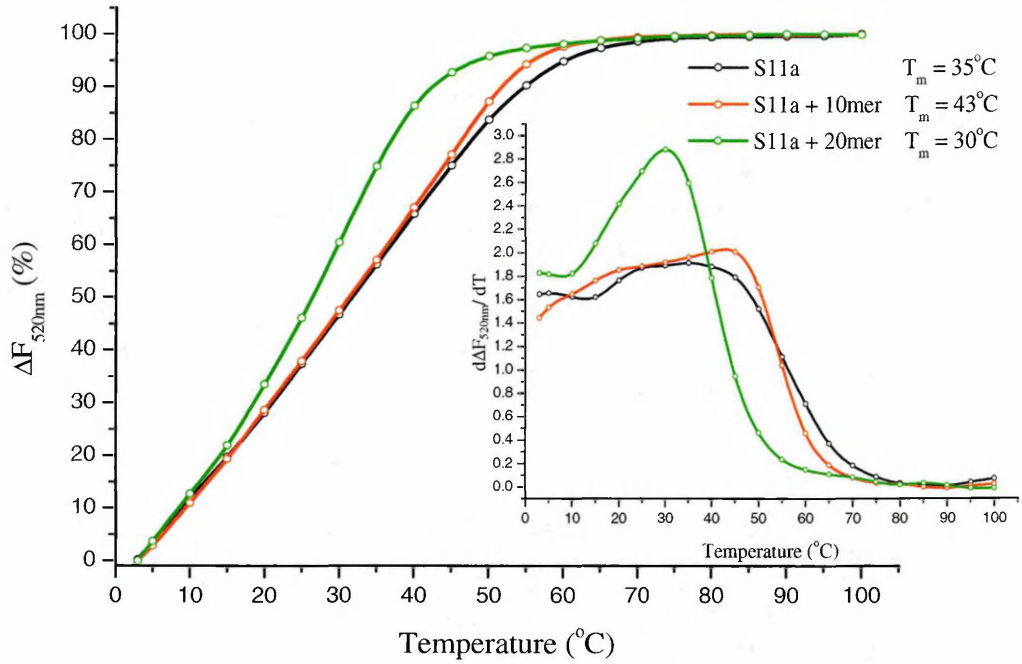


Figure 5.9. Thermal denaturation of S11a. Comparison of the fluorescence melting profiles of S11a with the 10 mer and 20 mer peptide.

Figure 5.10 shows the fluorescence melting profile of the truncated aptamers (S11b and S11c) with the MUC1/Y peptides. The addition of the 10 mer peptide to S11b causes a change in the melting profile as two T_m are observed (7°C and 20°C) although the there is no significant difference in the later T_m from that of free S11b (21°C), whilst the 20 mer peptide reduces the T_m to 16°C (figure 5.10A). The addition of the 10 mer and the 20 mer peptides to S11c causes the melting profile to change to provide a single T_m of 26°C and 16°C respectively from S11c (20°C and 45°C). If the later T_m of S11c (45°C) is considered, then the addition of the 10mer peptide has decreased thermostability of the aptamer by *ca.* 20°C, whilst the 20 mer has caused a 30°C decrease (figure 5.10B). Thus, the MUC1/Y peptides have the greatest effect on the melting temperature of the S11c aptamer compared to the S11a and S11b.

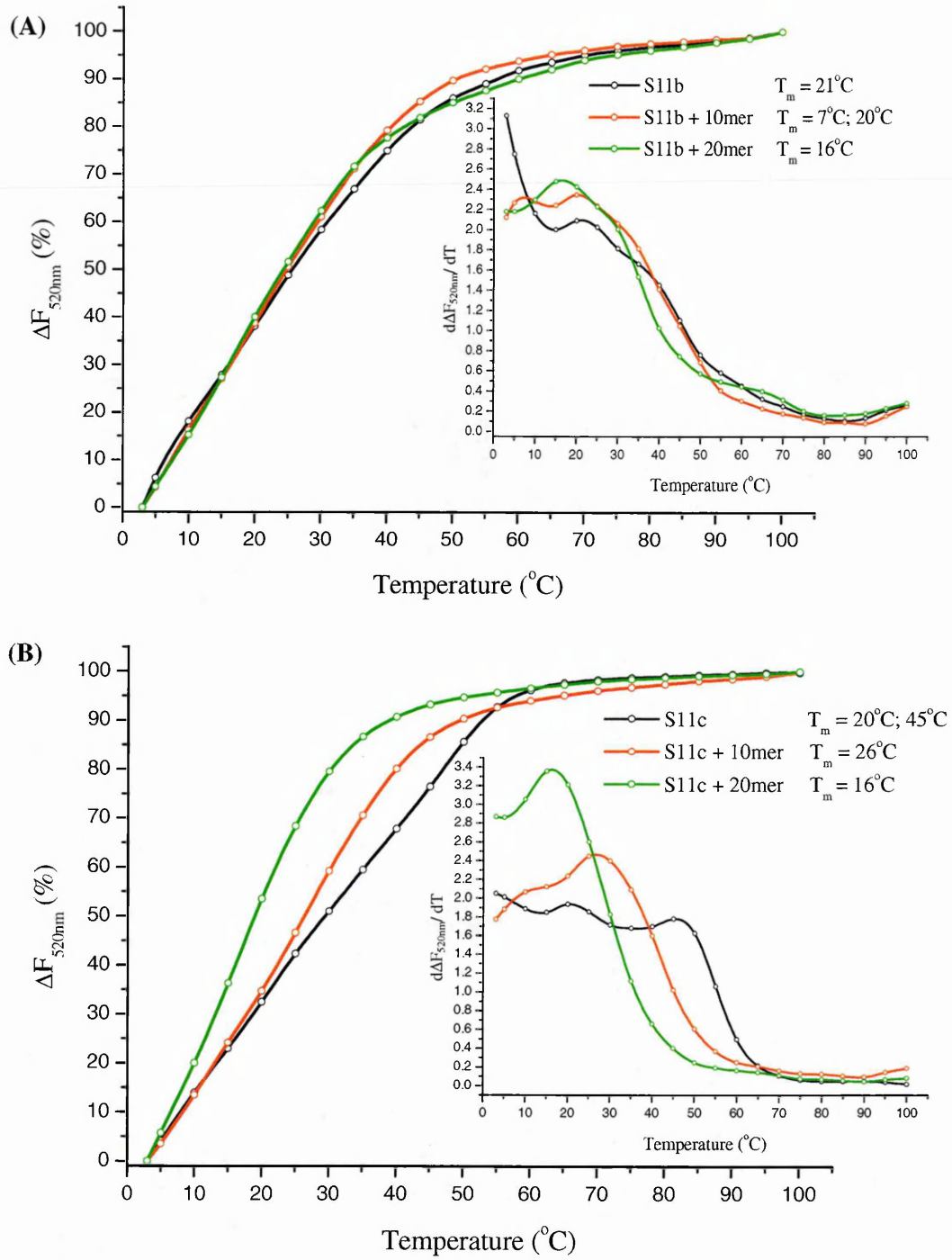


Figure 5.10. Thermal denaturation of S11b and S11c. (**A**) Comparison of the fluorescence melting profiles of S11b with the 10 mer and 20 mer peptide. (**B**) Comparison of the fluorescence melting profiles of S11c with the 10 mer and 20 mer peptide.

The thermal denaturation of the scrambled versions of the S11b and S11c (S11b-smd and S11c-smd respectively) used as controls with the MUC1/Y peptides is shown in figure 5.11. The S11b-smd has very similar T_m values as S11b. Addition of the 10 mer produces a second T_m (20°C and 35°C) and the 20 mer decreases the T_m to 16°C from 21°C (S11b-smd free) (figure 5.11A). Contrary to S11c, the addition of the

MUC1/Y peptides to S11c-smd has less of an effect on the T_m, where the 10 mer and 20 mer decreased the T_m to 42°C and 36°C from an original T_m of 47°C (figure 5.11B). The S11c-smd control aptamer appears to be more thermostable than the S11c aptamer.

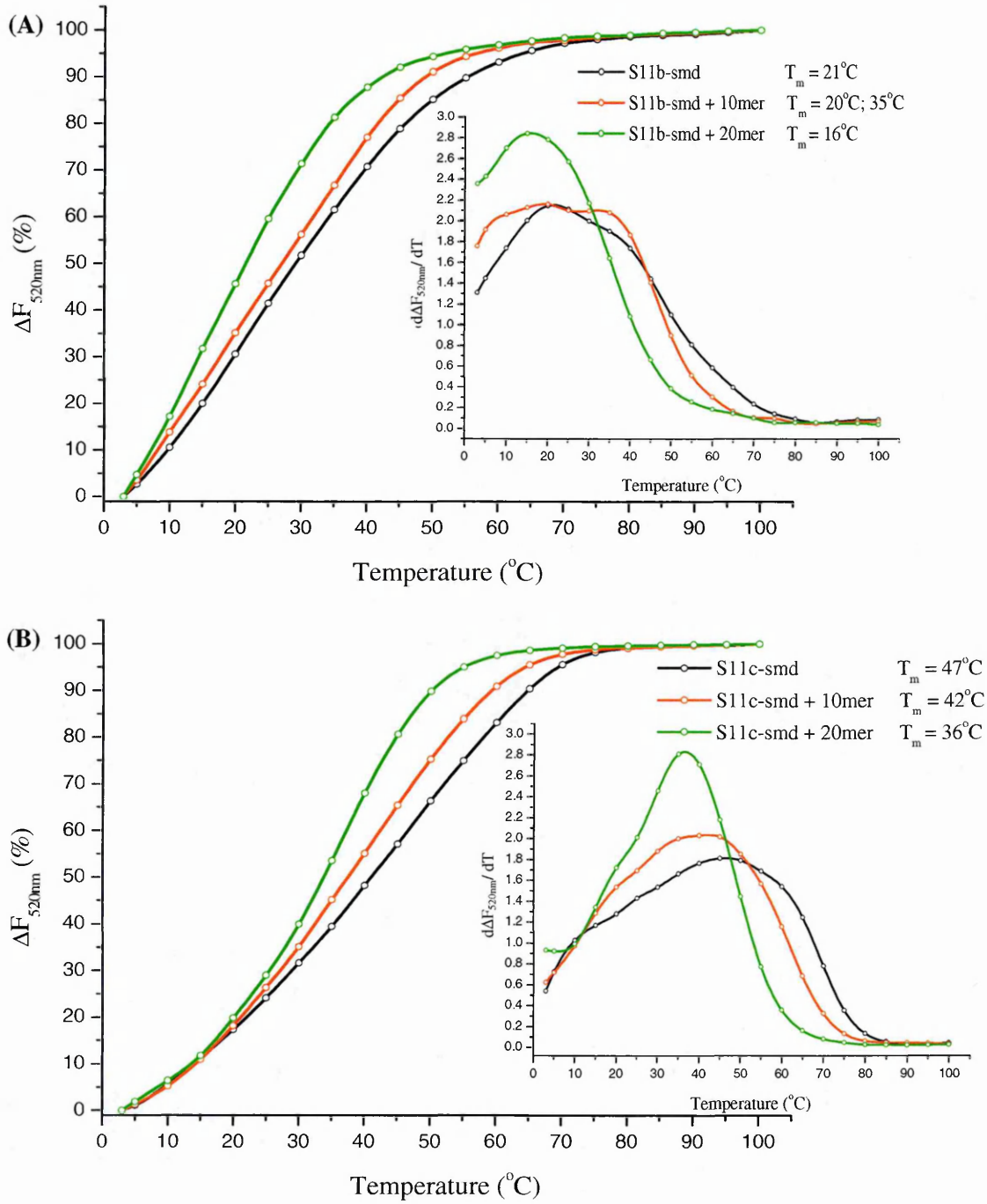


Figure 5.11. Thermal denaturation of S11b-smd and S11c-smd. (**A**) Comparison of the fluorescence melting profiles of S11b-smd with the 10 mer and 20 mer peptide. (**B**) Comparison of the fluorescence melting profiles of S11c-smd with the 10 mer and 20 mer peptide.

Although the pegylated aptamers S11b-PEG and S11c-PEG have shown increased thermostability compared to the unmodified aptamers, the thermostability is not extensively retained upon the addition of the MUC1/Y aptamers. Comparable T_m values of S11b-PEG (23°C and 19°C), shown in figure 5.12A, with S11b (20°C and 16°C) and likewise of S11c-PEG (35°C and 19°C), shown in figure 5.12B, with S11c (26°C and 16°C) upon the addition of the 10 mer and 20 mer peptide respectively, suggests that the pegylation of the aptamers has not affected the ability of the aptamers to bind to the MUC1/Y peptides.

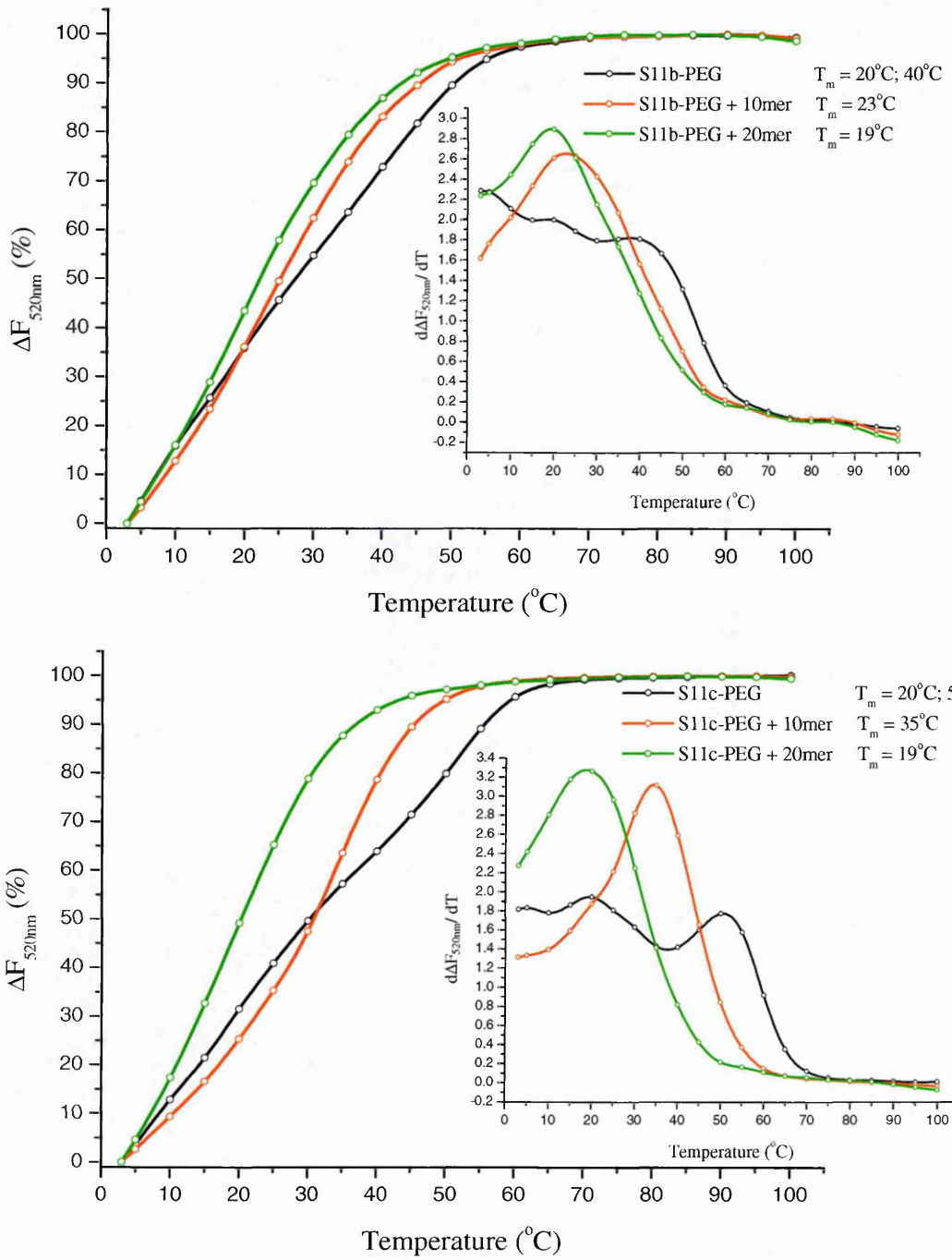


Figure 5.12. Thermal denaturation of S11b-PEG and S11c-PEG. (**A**) Comparison of the fluorescence melting profiles of S11b-PEG with the 10 mer and 20 mer peptide. (**B**) Comparison of the fluorescence melting profiles of S11c-PEG with the 10 mer and 20 mer peptide.

Plots of the S51a series with the control aptamers (S2.2 (MUC1) and SP68) without the addition of peptides is shown in figure 5.13 and 5.14 respectively, for comparing the effect of truncation on aptamers with the full length aptamer and the effect of pegylation with the unmodified versions. The S51a full length aptamer has two T_m (30°C and 52°C) and the truncation of the aptamer to S51b led to the loss of the higher T_m and decreased the lower T_m to 18°C. The T_m of S51b-PEG is comparable to that of the unmodified aptamer, suggesting that the attachment of the large polymer has not effected the primary conformation of the aptamer. The MUC1 control aptamer has a lower thermostability than the S51a by 12°C in both T_m (18°C and 40°C) (figure 5.13).

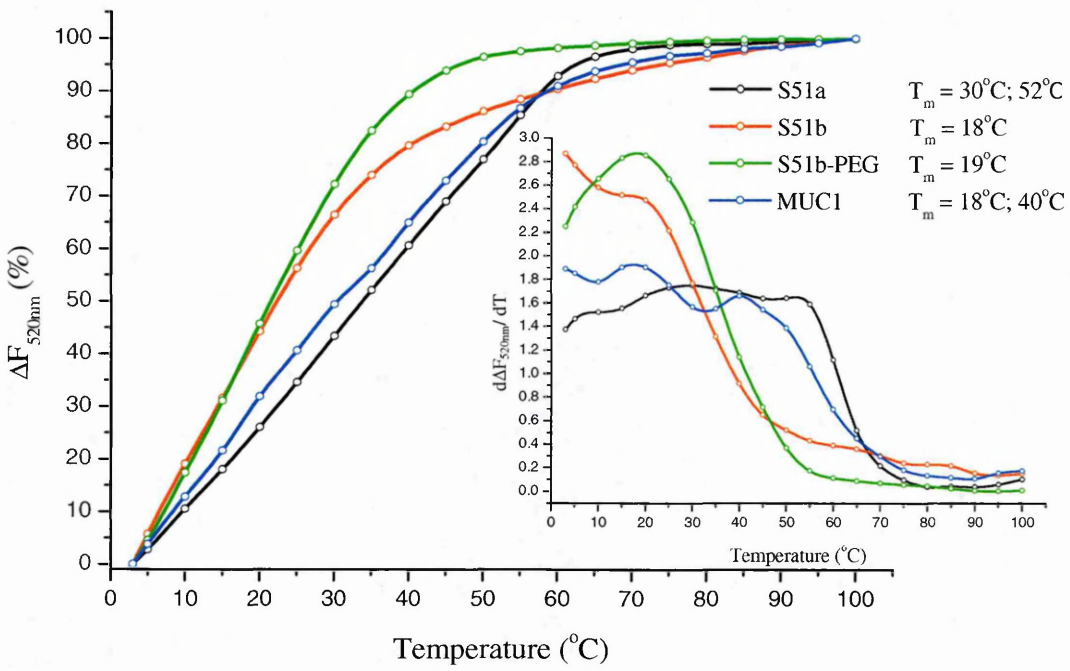


Figure 5.13. Thermal denaturation of S51 series. Fluorescence melting profiles of S51a, the truncated S51b versions and the control MUC1 aptamer.

The truncated loop region (S51c) of S51a has maintained the two T_m though with a 12°C reduction in the first T_m to 18°C. The T_m of S51c-PEG (20°C and 61°C) is slightly higher than the unmodified aptamer and moreover, the second T_m of S51c-PEG at 61°C is also higher than S51a at 52°C. The SP68 control aptamer has a single T_m of 25°C, overall relatively lower than the tm of the S51c aptamers, although more comparable with the S11b and S51b aptamers (figure 5.14).

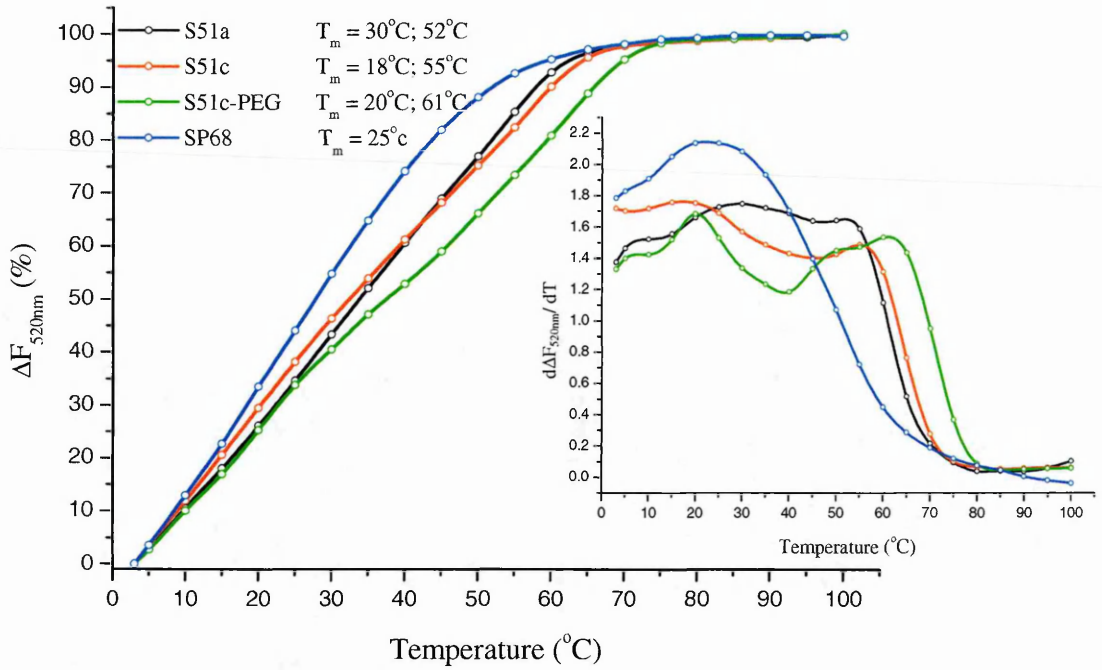


Figure 5.14. Thermal denaturation of S51 series. Fluorescence melting profiles of S51a, the truncated S51c versions and the control SP68 aptamer.

The effect of the MUC1/Y peptides added to the S51a aptamer is shown in figure 5.15. The addition of the 10mer peptide retains the two T_m values of S51a (30°C and 52°C), although the thermostability is reduced by 10°C and 16°C respectively. The addition of the 20mer peptide appears to affect the conformation, whereby only a single T_m is observed at 29°C, comparable to the first T_m of 30°C.

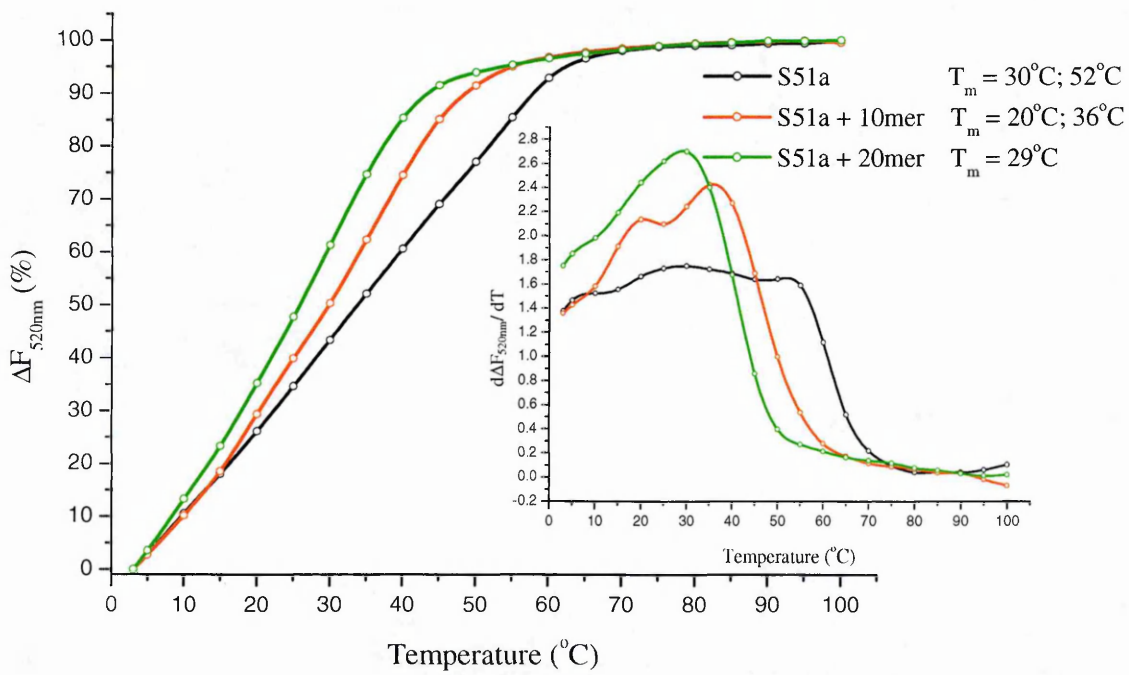


Figure 5.15. Thermal denaturation of S51a. Comparison of the fluorescence melting profiles of S51a with the 10 mer and 20 mer peptide.

Figure 5.16 shows the fluorescence melting profile of S51b (variable region) and S51c (loop region) with the MUC1/Y peptides. This truncated aptamer, S51b has a single T_m (18°C) that is significantly lower than that of the first T_m of S51a (30°C). The addition of the 10 mer and 20 mer peptide increases the thermostability slightly to 25°C and 21°C respectively (figure 5.16A). The S51c (loop region) retains the two T_m (18°C and 55°C) as present in S51a (30°C and 52°C). However, the first T_m is reduced by 12°C. Conversely, upon the addition of the 10 mer and 20 mer peptides, only a single T_m is observed at 37°C and 30°C respectively, thus appearing that the primary T_m value of S51c (18°C) has increased by 19°C and 12°C respectively (figure 5.16B).

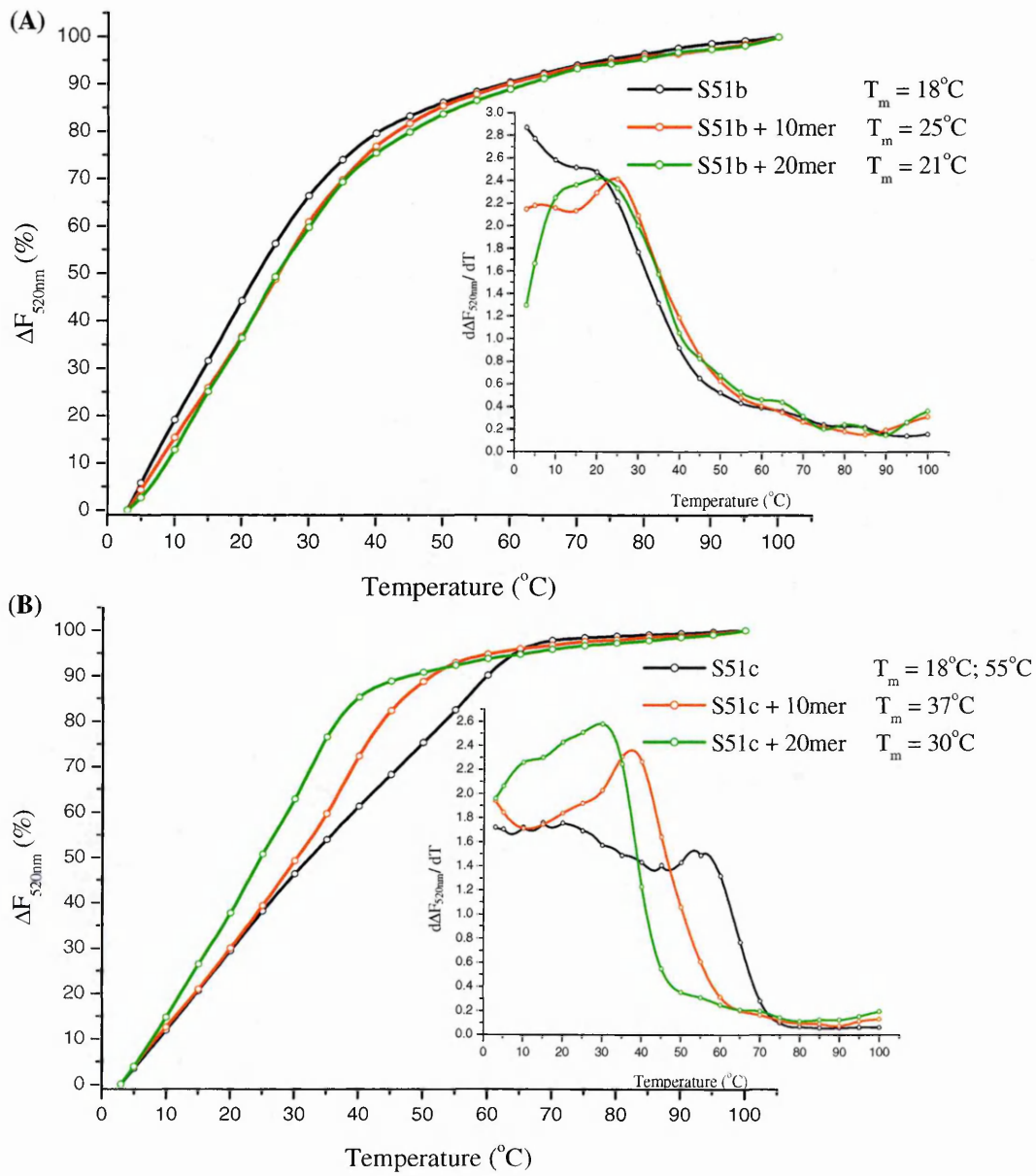


Figure 5.16. Thermal denaturation of S51b and S51c. (**A**) Comparison of the fluorescence melting profiles of S51b with the 10 mer and 20 mer peptide. (**B**) Comparison of the fluorescence melting profiles of S51c with the 10 mer and 20 mer peptides.

The S51b-PEG shows comparable T_m (19°C) values with the unmodified S51b (18°C) even with the addition of the 20 mer peptide resulting in a T_m of 23°C (S51b-PEG) compared to 25°C (S51b). The S51b-PEG (33°C) has an 8°C higher stability with the 10 mer peptide than the S51b (25°C) (figure 5.17).

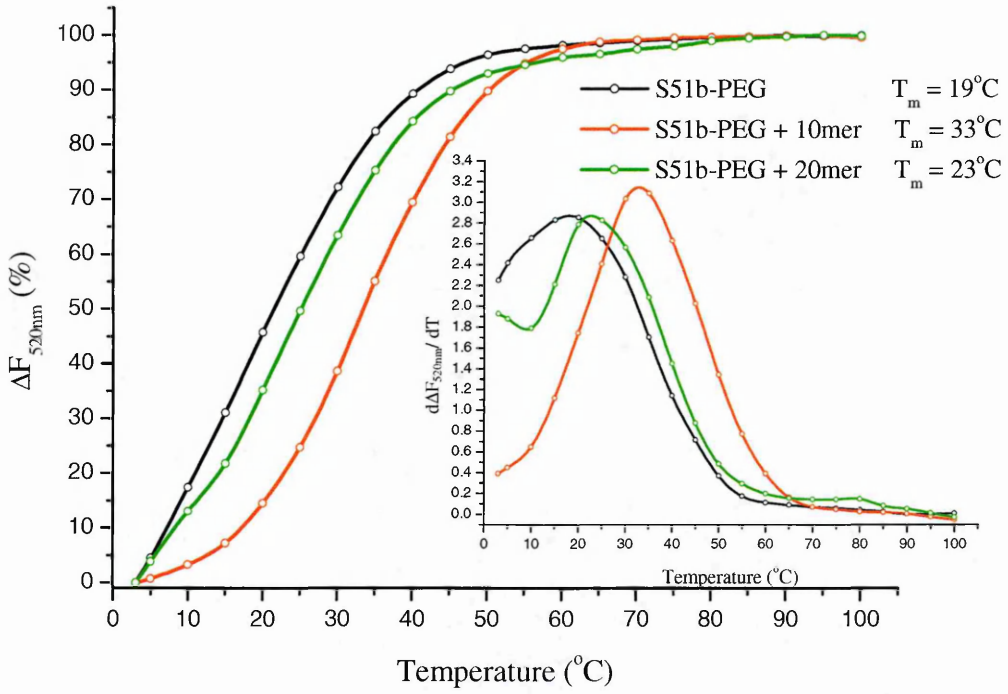


Figure 5.17. Thermal denaturation of S51b-PEG. Comparison of the fluorescence melting profiles of S51b-PEG with the 10 mer and 20 mer peptides.

S51c-PEG has been able to maintain the two T_m (20°C and 61°C) of the unmodified S51c (18°C and 55°C), even with the addition of the 10 mer peptide (36°C and 48°C) and 20mer peptide (20°C and 50°C). The S51c-PEG has slightly increased thermostability compared to the S51c and, although the stability is reduced to some extent with the addition of the MUC1/Y peptides, the Tm values are still comparable to that of the unmodified S51c (figure 5.18).

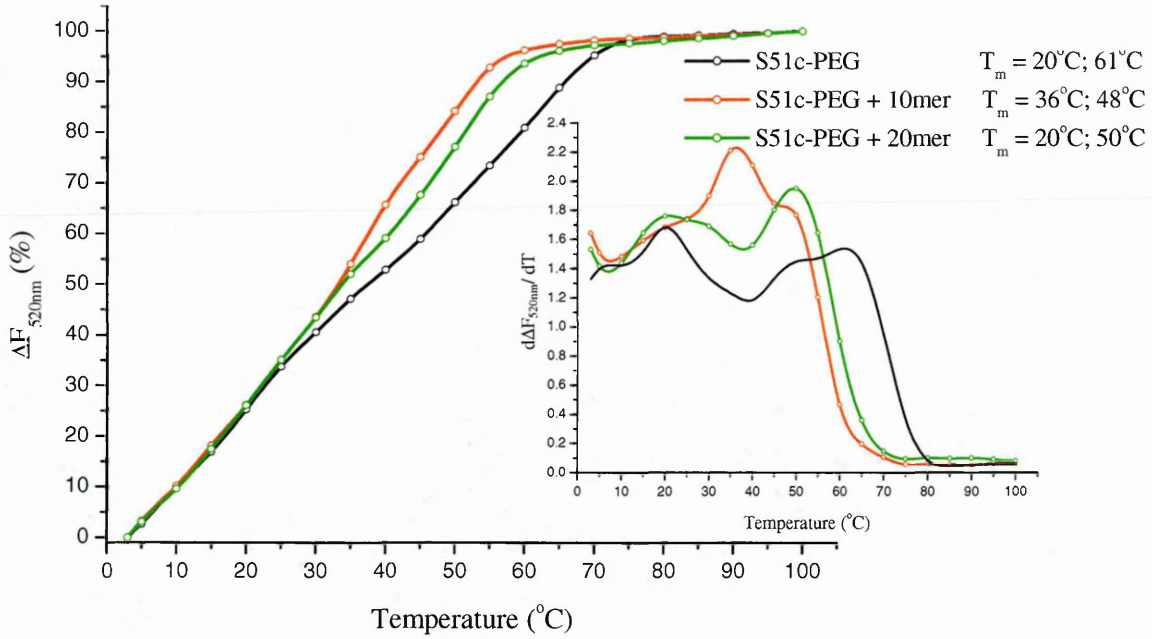


Figure 5.18. Thermal denaturation of S51c-PEG. Comparison of the fluorescence melting profiles of S51c-PEG with the 10 mer and 20 mer peptides.

The MUC1 aptamer has two T_m (18°C and 40°C) which goes to a single T_m on the addition of the 10 mer (29°C) and 20 mer (22°C) peptides. This trend of initially two T_m values and changing to only one T_m was also observed with the S11c (20°C and 40°C); 10 mer (26°C) and 20 mer (16°C) and with S51c (18°C and 55°C); 10 mer (37°C) and 20 mer (30°C) (figure 5.19).

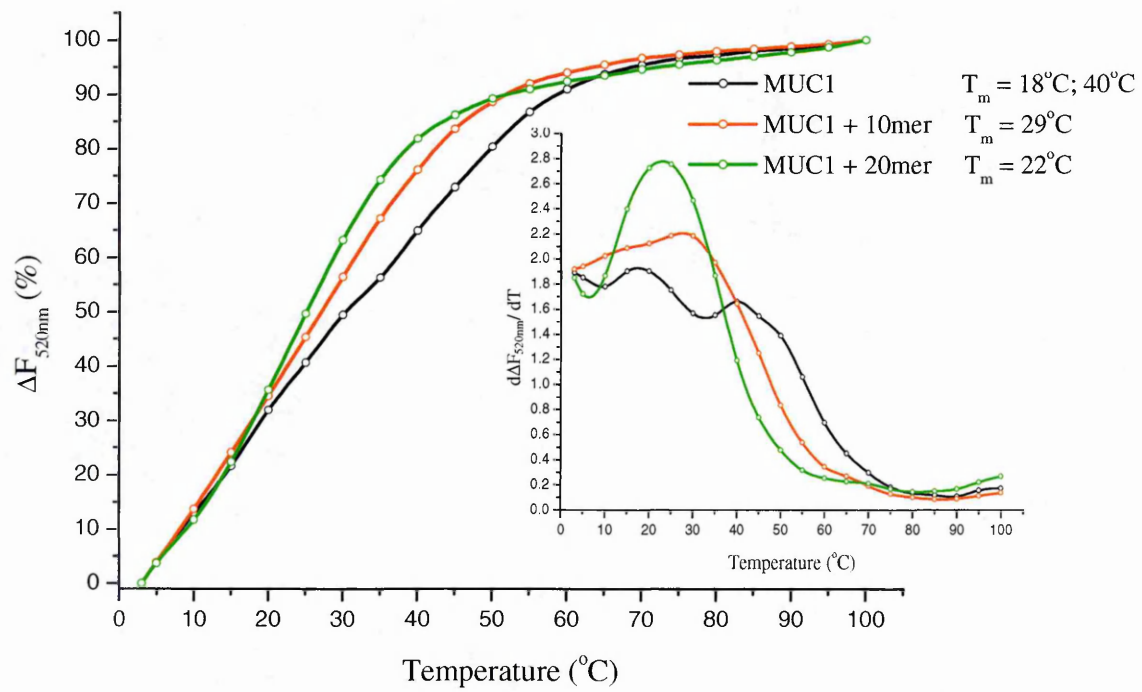


Fig 5.19 Thermal denaturation of MUC1. Comparison of the fluorescence melting profiles of MUC1 with the 10 mer and 20 mer peptides.

The SP68 has a T_m of 25°C, whereby significant changes ate not observed upon the addition of the 10 mer peptide (27°C) or the 20 mer peptide (29°C). SP68 has mainly retained its thermostability upon the addition of the MUC1/Y peptides (figure 5.20).

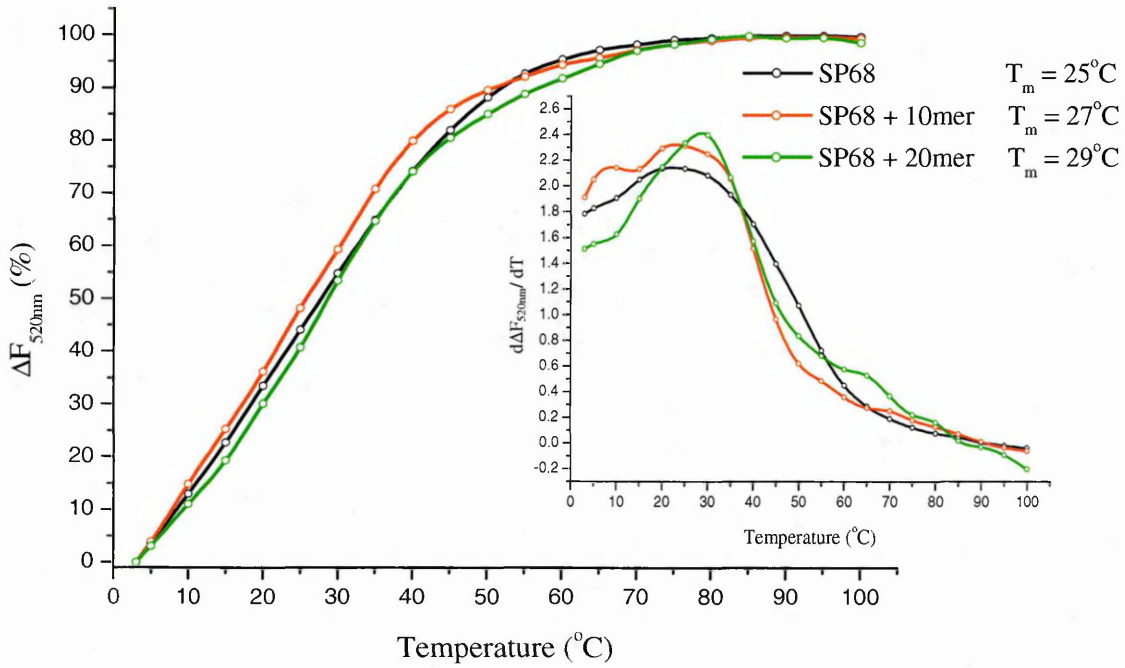


Figure 5.20. Thermal denaturation of SP68. Comparison of the fluorescence melting profiles of SP68 with the 10 mer and 20 mer peptides.

Plots of the S75a series without the addition of the MUC1/Y peptides is shown in figure 5.20. The full length S75a has three T_m (20°C, 35°C and 59°C) and the S75b (variable region) has a single T_m of 19°C comparable to the first T_m of S75a. The control aptamer, S75b-smd has two T_m (20°C and 47°C) of which the first T_m value is the same as the first T_m of S75a, and the second T_m (47°C) is between the second and third T_m values of S75a. The T_m of S75b-PEG (20°C and 40°C) is more comparable with the values of S75b-smd than the unmodified S75b, whereby a second T_m at 40°C is introduced (figure 5.20A). S75c (loop region) has T_m of 22°C, 43°C and 63°C, retaining the three T_m values of the full length S75a with slightly greater thermostability. The S75c-PEG has similar T_m values (21°C, 35°C and 63°C) to the unpegylated S75c with the exception of the second T_m values which is lower by 8°C (figure 5.21B).

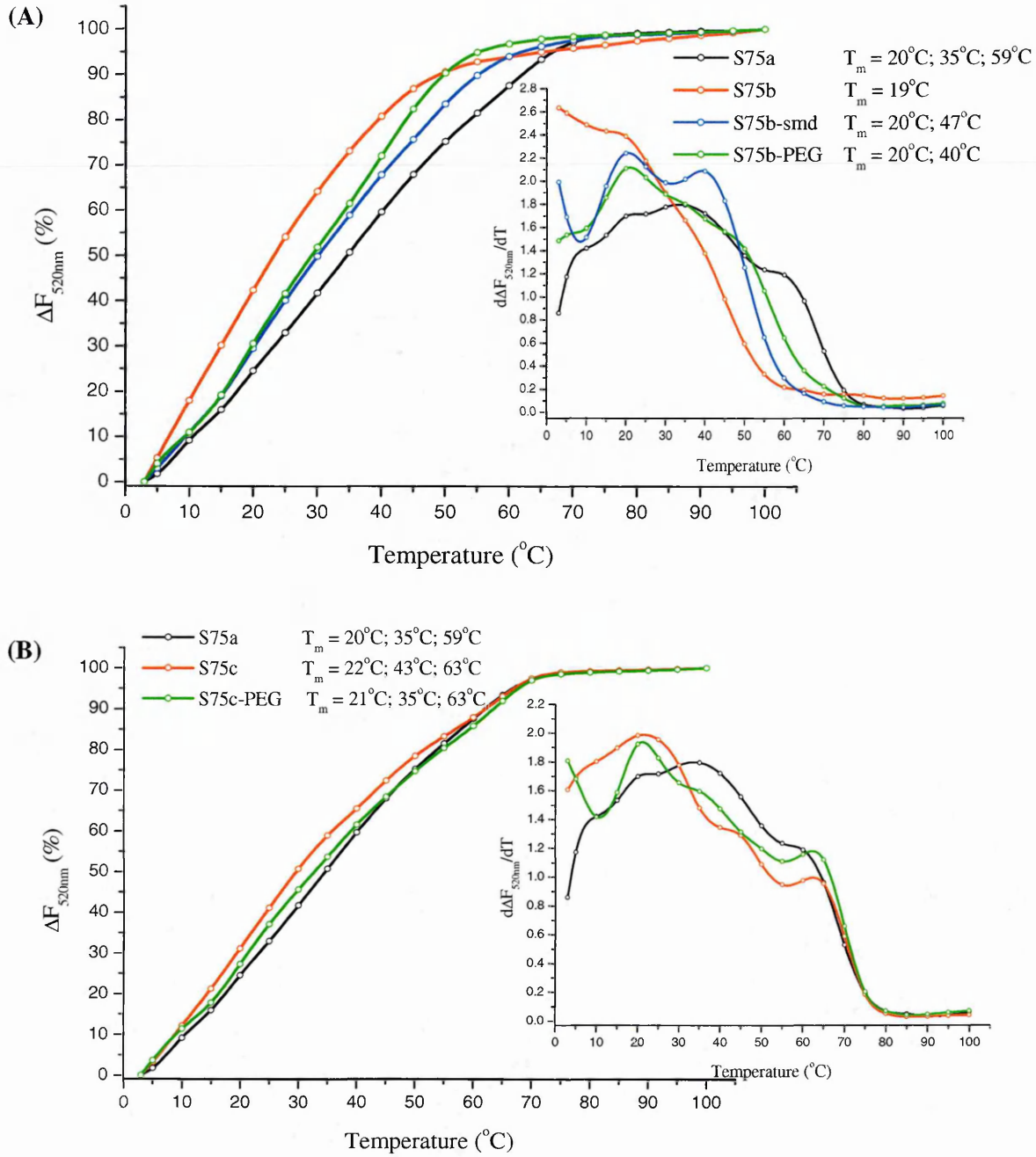


Figure 5.21. Thermal denaturation of S75 series. (**A**) Fluorescence melting profiles of S75a with the S57b versions (**B**) Fluorescence melting profiles of S75a with the S75c versions.

The effect of the MUC1/Y peptides added to the S75a aptamer is shown in figure 5.22. The addition of the 10 mer peptide appears have caused a loss in the third T_m at 59°C and exhibit to T_m of 20°C and 43°C, whilst the addition of the 20 mer peptide only displays a single T_m of 28°C, 8°C higher than the first T_m of the free S75a and with the 10 mer peptide.

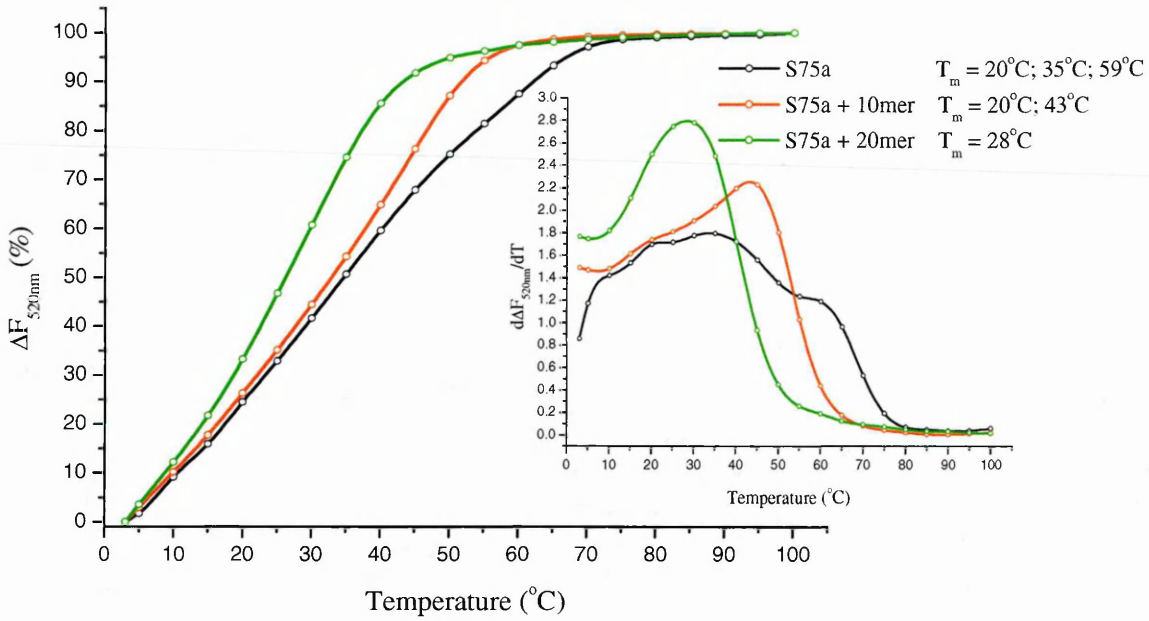


Figure 5.22. Thermal denaturation of S75a. Comparison of the fluorescence melting profiles of S75a with the 10 mer and 20 mer peptide.

Figure 5.23 shows the fluorescence melting profile of S75b (variable region) and S75c (loop region) with the MUC1/Y peptides. The thermostability of S75b remains unchanged compared to the free aptamer upon the addition of the 10mer peptide at a T_m of 19°C. However, the addition of the 20 mer peptide increases the thermostability by 4°C to 23°C (figure 5.23A). The addition of the MUC1/Y peptides appears to reduce the thermostability and change the conformation of S75c, as only two T_m (20°C and 30°C) are observed with the addition of the 10 mer peptide and only a single T_m (22°C) with the 20 mer peptide compared to the T_m values of the free S75c (22°C, 43°C and 63°C) (figure 5.23B).

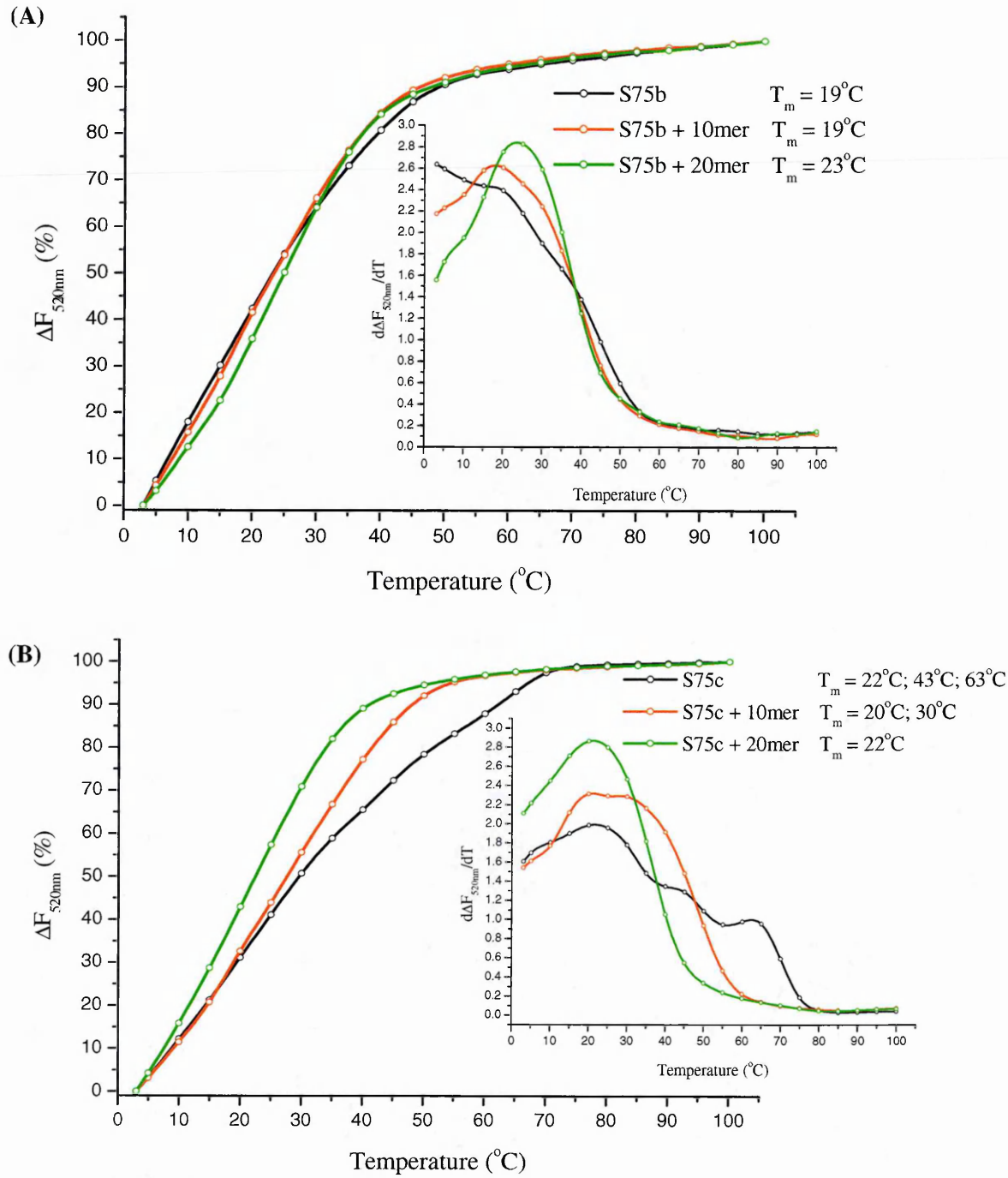


Figure 5.23. Thermal denaturation of S75b and S75c. (**A**) Comparison of the fluorescence melting profiles of S75b with the 10 mer and 20 mer peptide. (**B**) Comparison of the fluorescence melting profiles of S75c with the 10 mer and 20 mer peptides.

The control aptamer S75b-smd which is the scrambled version of the S75b has two T_m (20°C and 47°C) as opposed to single T_m like the S75b (19°C), although upon the addition of the MUC1/Y peptides, S75b-smd only remains with a single T_m of 22°C and 24°C for the 10 mer and 20 mer peptide respectively. The

T_m values following the addition of the MUC1/Y peptides are comparable with the T_m of S75b at 19°C and 23°C upon the addition of the 10 mer and 20 mer peptide respectively (figure 5.24.).

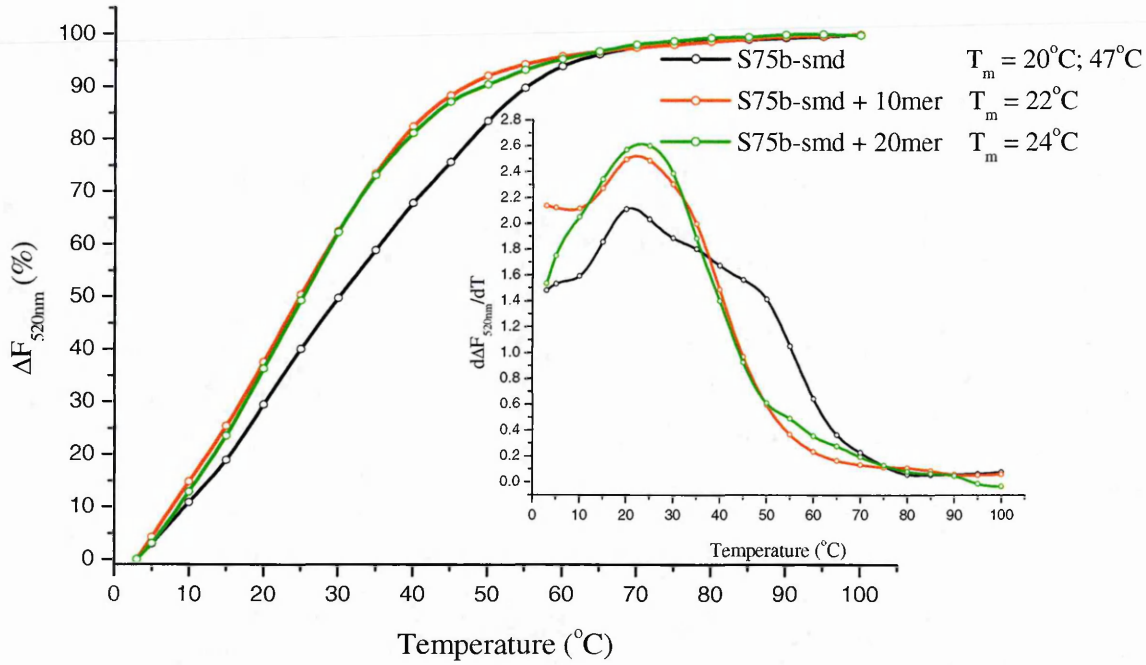


Figure 5.24. Thermal denaturation of S75b-smd. Comparison of the fluorescence melting profiles of S75b-smd with the 10 mer and 20 mer peptide.

The S75b-PEG has two T_m (20°C and 40°C) of which the first T_m is comparable to the single T_m of unmodified S75b (19°C). The addition of the 10 mer peptide still retains the two T_m (19°C and 34°C) but with slightly reduced thermostability, whilst the addition of the 20mer peptide changes the conformation of S75b-PEG resulting in a single T_m of (28°C) (figure 5.25A). The T_m values of S75c-PEG (21°C, 35°C and 63°C) are similar to the unmodified S75c aptamer (22°C, 43°C and 63°C) with the exception that the second T_m is lowered by 8°C. The addition of the MUC1/Y peptides appears to bring about a conformational change in the aptamer, whereby only a single Tm at 39°C and 22°C for the 10 mer and 20 mer peptides respectively is exhibited. Overall, the addition of the peptide reduces the thermostability of the original aptamer conformation (figure 5.25B).

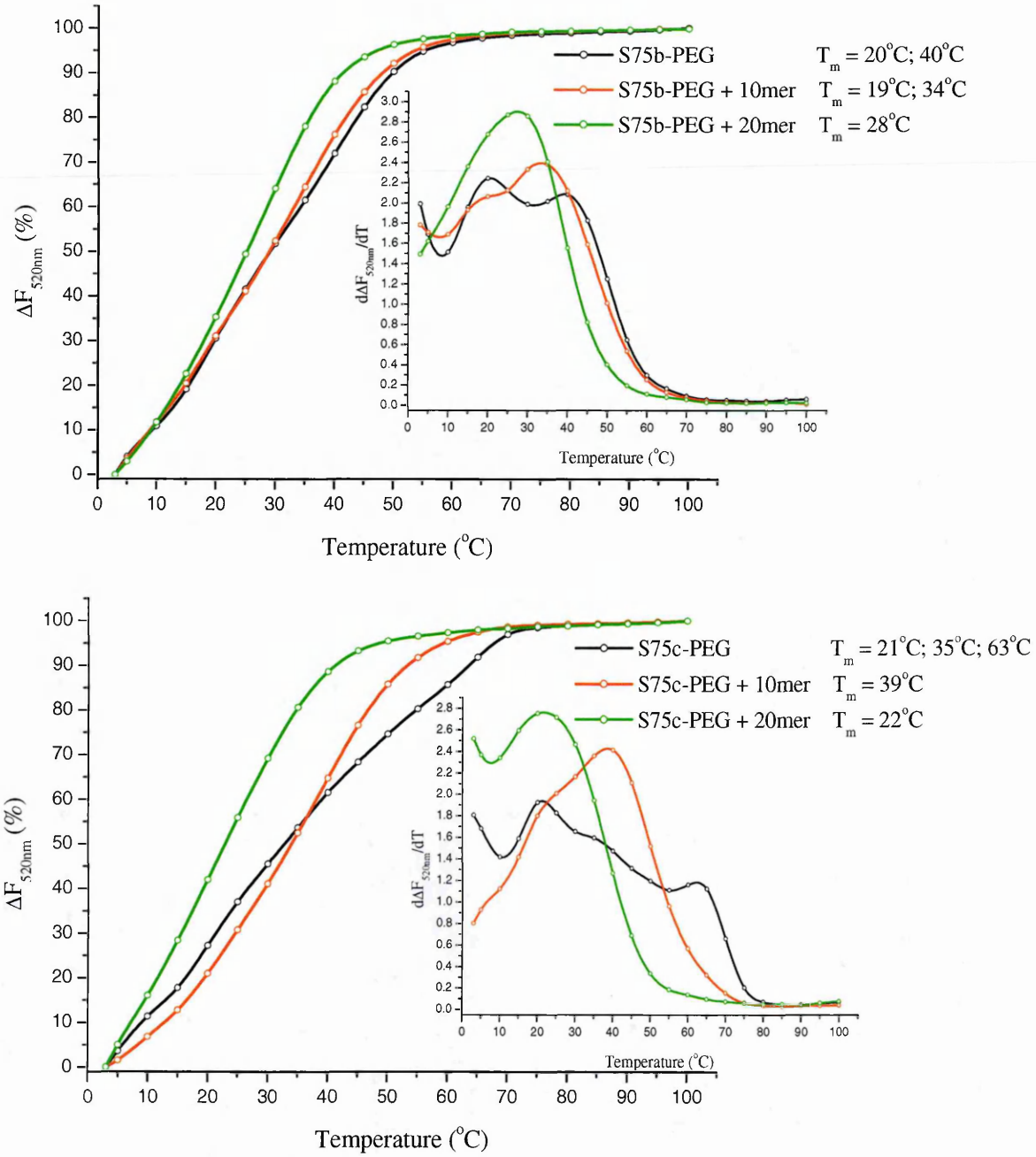


Figure 5.25. Thermal denaturation of S75b-PEG. Comparison of the fluorescence melting profiles of S75b-PEG with the 10 mer and 20 mer peptide.

This thermal DNA denaturation technique has been an insightful method to ascertain the potential conformational changes that may occur at particular temperatures, upon binding to the peptides and from any structural modifications. From the three different series of aptamers, including the control aptamers, single, double and triple melting points for an aptamer were observed on occasion. This could be attributed to a number of factors which includes the possibility that there is a mixture of conformations of the aptamer in solution. Alternatively, there maybe just one conformation which contains one, two or

three short stems which have different T_m values. The fluorescence melting curves obtained from all the aptamers has evidently shown that once the aptamer denatures in the presence of the peptide, re-folding of the aptamer back to its natural conformation was inhibited. In the case of the MUC1/Y peptides, the length of the peptide played a significant role as the 20 mer peptide was able to denature the aptamer to a greater extent than the 10 mer peptide prior to the first TDC. Thus, greater reduced thermostability of the aptamer was observed with the 20 mer peptide than the 10 mer peptide during the first TDC (table 5.0). This was exhibited by the progressive decrease in the fluorescence of the SYBR Green dye through each TDC. As this trend was also observed with the control aptamers, this could be a possible indication, that as the scrambled and selected aptamers possess identical base composition and charge, once unstructured, can interact similarly with the peptides. In the case of all three aptamer series, the truncated b versions (central variable region) of the full length aptamer presented reduced T_m, whilst the truncated c versions (stem-loop region) showed greater stability than the b version and comparable or higher T_m values to the full length aptamers. This is expected, as the truncated c versions contained the main structural features of the aptamers and their truncations were based on these features. The scrambled control aptamer S11b-smd had the same T_m as S11b, whilst S11c-smd had greater T_m than the full length S11a. S75b-smd showed greater thermostability than the S75b. All the pegylated aptamers exhibited greater thermostability than the unmodified aptamers except for S51b-PEG which had a similar T_m to the unmodified S51b. This typically indicates that the pegylation of the aptamers aids in improving the thermostability of the aptamers. The pegylated aptamers retain their increased thermostability compared to the unmodified aptamers upon the addition of the MUC1/Y peptides. In addition, all the pegylated aptamers with the exception of the S51b-PEG possess two or more melting temperatures, of which one is greater than 37°C. This implies that at physiological temperature, the pegylation of the aptamers would assist in the maintenance of the natural aptamer conformation. Furthermore, upon the pegylated aptamer binding to the target peptides the aptamer is still able to adopt an open structure, further strengthening the affinity of the aptamer for the peptides. Thus, pegylated aptamers would be the most applicable when considered for studies involving physiological temperatures as initially the original structure of the aptamer is stabilised and the large polymer also does not demonstrate any form of interference in the binding of the aptamers with the MUC1/Y peptides. The control non-MUC1/Y aptamers (MUC1 and

SP68) shows comparable T_m values to the MUC1/Y aptamers, however, SP68 appears to be the only aptamer which shows least effect on the conformation of the aptamer upon the addition of the peptides. This is an indication that SP68 may not have a significant affinity and specificity for the MUC1/Y peptides and should be expected, as this aptamer has been selected for a non-MUC1 related target (table 5.0).

APTAMER	MELTING TEMPERATURE		
	Free Aptamer	Aptamer + 10 mer	Aptamer + 20 mer
S11a	35°C	43°C	30°C
S11b	21°C	7°C; 20°C	16°C
S11b-smd	21°C	20°C; 35°C	16°C
S11b-PEG	20°C; 40°C	23°C	19°C
S11c	20°C; 45°C	26°C	16°C
S11c-smd	47°C	42°C	36°C
S11c-PEG	20°C; 51°C	35°C	19°C
S51a	30°C; 52°C	20°C; 36°C	29°C
S51b	18°C	25°C	21°C
S51b-PEG	19°C	33°C	23°C
S51c	18°C; 55°C	37°C	30°C
S51c-PEG	20°C; 61°C	36°C; 48°C	20°C; 50°C
S75a	20°C; 35°C; 59°C	20°C; 43°C	28°C
S75b	19°C	19°C	23°C
S75b-smd	20°C; 47°C	22°C	24°C
S75b-PEG	20°C; 40°C	19°C; 34°C	28°C
S75c	22°C; 43°C; 63°C	20°C; 30°C	22°C
S75c-PEG	21°C; 35°C; 63°C	39°C	22°C
MUC1	18°C; 40°C	29°C	22°C
SP68	25°C	27°C	29°C

Table 5.0. Melting temperatures of the MUC1/Y aptamers. Melting temperature of the first TDC is shown for each aptamer before the addition of the peptides and with the each peptide.

5.7 CONCLUSION

The three qualitative characterisation techniques (EMSA, affinity column chromatography and DNA thermal denaturation) performed have revealed numerous aspects of the MUC1/Y aptamers behaviour towards their target peptides. Whilst the EMSA at first hand does not show any obvious complex formation of the aptamer with the peptides, the EMSA results do indicate for all the aptamers analysed that upon heating of the aptamer and peptide together, even for a period as short as 5 minutes, the peptide

is capable of binding to the denatured aptamer and inhibiting the renaturation of the aptamer. On the EMSA analysis this was indicated by the disappearance of the aptamer band, as the EtBr was unable to intercalate with an open DNA structure. This was further emphasised by the fluorescence thermal denaturation studies, whereby a progressive decrease in the fluorescence was observed with each thermal denaturation cycle. Both the EMSA and the fluorescence thermal denaturation studies demonstrated that the conjugation of the aptamers to a large polymer PEG did not affect the binding of the aptamers to the peptides prior to denaturation and during denaturation.

The results obtained from the fluorescence thermal denaturation studies have provided greater clarification of the outcomes obtained from the affinity column chromatography studies. As an overall trend for all three series, it appears that the conformation the aptamer has prior to the interaction with the peptides is different to the conformation the aptamer adopts upon binding to the peptides and the latter potentially determines the affinity of the aptamer. As an example, upon binding to the 10 mer peptide, S11b has a T_m of 7°C and 20°C, S11b-smd has a T_m of 20°C and 35°C, and S11b-PEG has a T_m of 20°C and 40°C, with a 10 mer affinity column binding of 32 %, 27% and 20.5% respectively. This trend implies that as the second T_m of each aptamer increases, greater secondary structure of the aptamer is present and combined with the low affinity column binding results, this further suggests that a more open conformation of the aptamer yields greater affinity of the aptamer for the target peptide. However, this trend is not consistent with all the aptamers. As an example, upon binding to the 20 mer peptide, S11b has a T_m of 16°C, S11b-smd has a T_m of 16°C, and S11b-PEG has a T_m of 19°C, with a 20 mer affinity column binding of 21 %, 20% and 47% respectively. In this case S11b-PEG has a low T_m (19°C) yet the binding in the affinity column to the 20 mer peptide was 47%. This suggests that only an open conformation is not the main determinant for the binding of the aptamer to the peptides, but also the overall conformation of the aptamer plays a significant role in the binding and thus indicating structural specificity of the aptamers for their target upon binding. However, differences in aptamer T_m that are below room temperature are of no relevant significance with respect to the binding, as these would be mute point unless experiments were performed at 4°C, and would of course not be useful in the design of a therapeutic that needs to function at physiological temperatures. Thus, differences between S11b, for

example (T_m of 16°C) and S11b-PEG (T_m of 19°C) would not actually portray any specific binding, as both aptamers would be unstructured at a physiological temperature. On the other hand, the pegylated aptamer has been shown to have a higher affinity for the peptide. Thus, though the selection has been performed with different methodologies and consensus sequences have been obtained, which are clearly important for binding, it appears that the binding of the selected aptamers to the MUC1/Y peptides is not based on a key-lock mode of binding, where the aptamer retains a fixed secondary structure, but it is more based on the aptamers and peptides adopting complementary structures that lead to optimal binding. Furthermore, of the two control aptamers, MUC1 displayed comparable T_m values to the MUC1/Y aptamers whilst the SP68 exhibited the least change in conformation upon the addition of the peptides. Likewise, if however scrambled peptides (10 mer and 20 mer) were to be used, there is a high probability that the aptamers may interact with the scrambled peptides as with the unscrambled peptides. This could be a result of the aptamers being more structural specific to the peptides than sequence specific. Thus, if the structure of the scrambled peptides remained unaffected, then significant differences in the denaturation of the aptamers may not be observed.

Therefore to conclude the overall findings from the combination of all three techniques, it appears that the conformation which the aptamer adopts upon binding to the peptide then further determines its affinity for the target. Typically if the conformation is resulting in a more open structure, then the affinity for the target appears to be greater and the interaction between the aptamer and peptide can be irreversible.

CHAPTER SIX

IN VITRO CELL BINDING & INTERNALISATION STUDIES

6.0 INTRODUCTION

The qualitative analysis of series S11, S51 and S75 aptamers all indicated affinity for the MUC1/Y 10 mer and 20 mer peptides to varying degrees, depending upon the conformation they adopt, which is influenced by temperature. As the MUC1/Y peptide sequences are identical to the sequence present in the MUC1/Y protein [21]. expressed on many carcinoma cells lines, it was essential to determine whether the selected aptamers exhibit any form of binding to the full protein expressed on tumour cells. In addition, the aptamers possess different melting temperatures, and some aptamers were found to have T_m similar to that of the physiological temperature 37°C, whilst others have T_m values significantly higher or lower than 37°C. Therefore, it was also crucial to analyse whether there was a correlation in T_m value with the aptamer binding. To assess the binding of the aptamers, fluorescence activated cell sorting (FACs) technique was employed, which entails incubation of the fluorescently labelled aptamers with the cell lines of interest, followed by separation and analysis of these cells based on their size and fluorescence. Changes in the shift pattern of the histograms indicate whether binding of the aptamers to the MUC1/Y protein expressed on the cells has taken place when compared to the cytometric profile of the control cells (no aptamer). From the ten different cancer cell lines screened for MUC1 and MUC1/Y positive cell lines, three MUC1 positive MCF-7 (human breast carcinoma), DU145 (human prostate carcinoma) and CALU-6 (human lung carcinoma) and one MUC1 negative A498 (human kidney carcinoma) cell lines were used to conduct FACs analyses. Alongside the cell binding studies, cell internalisation of the aptamers on the same four cell lines was also investigated using fluorescence microscopy analysis.

6.1 APTAMER CELLULAR BINDING USING FACs

Each Cy3 fluorescently labelled aptamer of the S11 series (S11b, S11b-smd, S11b-PEG, S11c, S11c-smd, and S11c-PEG), the S51 series (S51b, S51b-PEG, S51c, and S51c-PEG), the S75 series (S75b, S75b-smd, S75b-PEG, S75c and S75c-PEG) and the control aptamer SP68 (selected against a non-MUC1 tumour target peptide) were incubated with MCF-7, DU145, CALU-6 and A498 cell lines in PBS at 37°C for 1 hour. Upon incubation, the cells were washed with PBS, centrifuged, resuspended in PBS and

analysed. Histograms of the S11 series analysed with the four carcinoma cell lines, clearly demonstrate that aptamer binding in the MUC1/Y positive cell lines (MCF-7, DU145 and CALU-6) has taken place, as shown by a shift in the cytometric profiles to the right of the control cell population (figure 6.0). MCF-7 cell line showed significant binding to S11b and S11c, but only marginally superior to that of their respective scrambled versions (S11b-smd and S11c-smd). S11b-PEG demonstrated equal or slightly greater binding than S11b, whilst S11c-PEG had much lower binding than S11c (figure 6.0, 1A-B). A similar binding trend was observed with the DU145 cell line but to a lesser extent than with the MCF-7 cell line. The S11b and S11c aptamers have comparable level of binding to their respective scrambled versions (S11b-smd and S11c-smd), whilst the S11b-PEG shows slightly greater binding than the unmodified S11b. On the other hand, the S11c-PEG has shown similar or slightly less binding than the S11c (figure 6.0, 2A-B). CALU-6 cell line has shown as strong binding as the MCF-7 cell line with the S11 aptamer series. The scrambled S11b-smd has a reduced affinity than S11b whereas the S11b-PEG has a similar binding profile to the unmodified aptamer. S11c-smd has an almost identical binding to the aptamer S11c and theS11c-PEG has lower binding than the unmodified aptamer. However, this decrease in binding is to a lesser extent to that observed with the MCF-7 cell line (figure 6.0, 3A-B). No binding of the S11 aptamer series was observed in the A498 MUC1/Y negative cell line with the exception of the S11c-smd aptamer where some binding maybe observed when compared to the other S11 aptamers (figure 6.0, 4A-B).

The FACs analyses of the S11 aptamer series revealed that these MUC1/Y selected aptamers have specificity for their target protein expressed on the carcinoma cells, as binding was observed with the MUC1/Y positive cell lines and not on the MUC1/Y negative cell line. Furthermore, the pegylation of the aptamers have shown to alter the conformation of the aptamer, which is marked by the difference in the T_m value from the unmodified aptamer. Thus, the alteration in the aptamer structure clearly correlates to the changes in the binding to the protein expressed by the cells, as in the case of S11b-PEG, where the increased thermal stabilisation of the aptamer from 45°C to 51°C leads to a reduction in binding. This is consistent with observations from the T_m study, where lower T_m values correlated with increased binding.

However, the increase in thermal stability of the S11b-PEG from 21°C to 40°C has led to no significant change in binding.

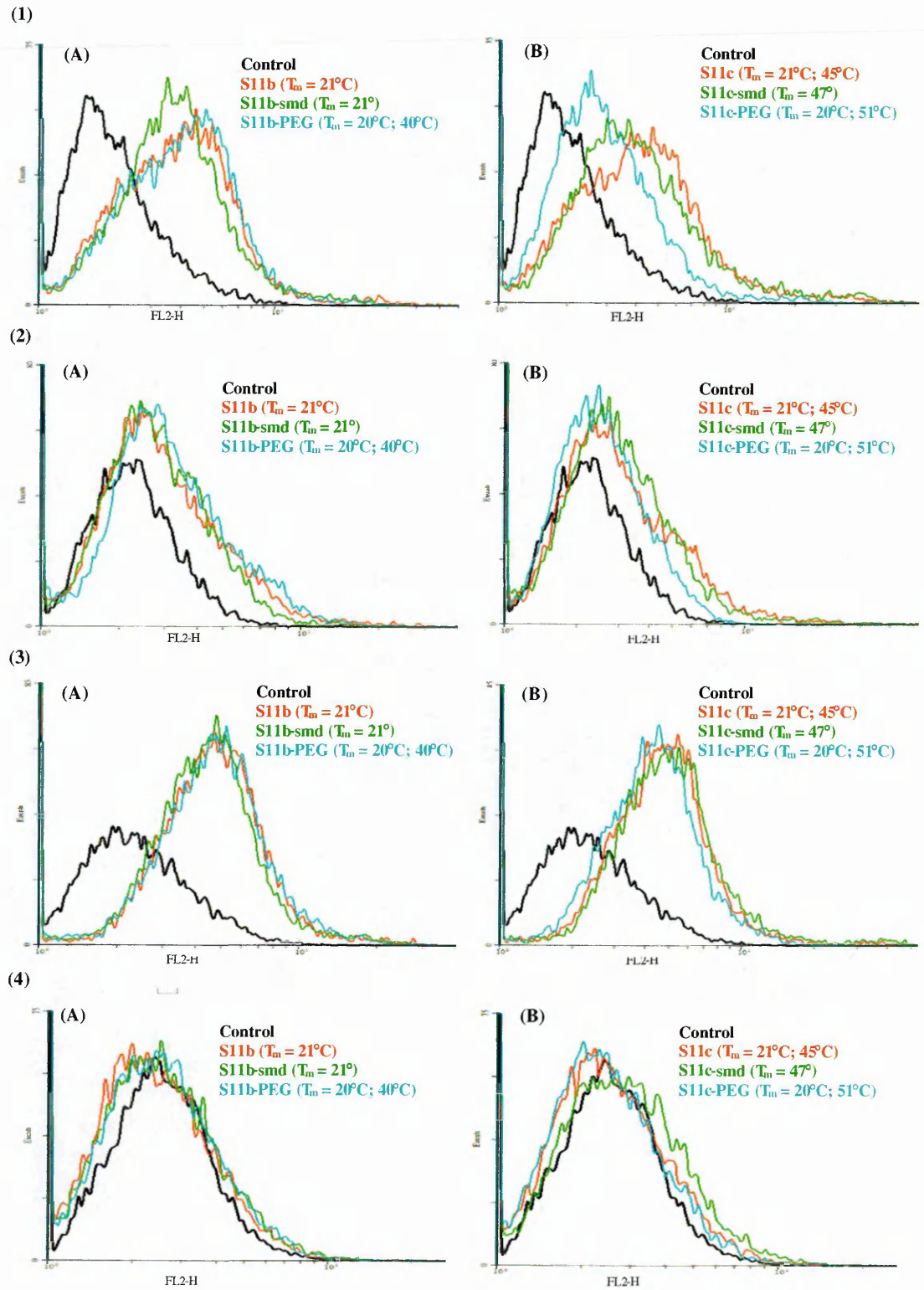


Figure 6.0. FACs histograms of S11 aptamer series. (**A**) S11b version (**B**) S11c version. (**1**) MCF-7 cell line, (**2**) DU145 cell line, (**3**) CALU-6 cell line and (**4**) A498 cell line.

Figure 6.1 shows the FACs analysis of the S51 series and the control aptamer SP68. Considerable level of binding has been observed of the S51 aptamer series with the MCF-7 cell line. The aptamers S51b-PEG and S51c-PEG have shown to retain their binding potential and are equal to their respective unmodified aptamers S51b and S51c. The control aptamer, SP68 has also shown binding and has a cytometric profile which is comparable to that of the S11b-PEG (figure 6.1, 1A-B), although this was expected as this aptamer was also selected against another tumour target (non-MUC1). The S51 aptamer series, however, has shown only minor binding to the DU145 cell lines, where again the S51b-PEG and S51c-PEG display comparable binding with the unmodified aptamers S51b and S51c respectively. Furthermore, the control aptamer SP68 has similar binding to the S51b aptamer (figure 6.1, 2A-B), as expected. The CALU-6 cell line showed variability in the binding to the different aptamers of the S51 series. S51b-PEG and S51c-PEG showed lower binding when compared to the unmodified aptamers S51b and S51c respectively. Moreover, although SP68 has shown binding, it is to a less extent than the S51b aptamer (figure 6.1, 3A-B). This series of aptamers have shown negligible level of binding and the shift like appearance of the cytometric profiles are likely to be attributed to the increase in the cell population and thus explaining the broadness of the peak. Additionally, the SP68 does not show any significant binding (figure 6.1, 4A-B).

Analysis of the S51 aptamer series has demonstrated that the greatest binding of the pegylated and unpegylated aptamers was observed with the MCF-7 cell line. The modified aptamers S51b-PEG (T_m: 19°C) and S51c-PEG (T_m: 20°C; 61°C) share similar T_m values with their respective unmodified aptamers S51b (T_m: 18°C) and S51c (T_m: 18°C; 55°C) along with exhibiting similar binding profiles with the DU145 cell line. Contrary to this, the aptamers S51b-PEG and S51c-PEG exhibit reduced binding than the unmodified aptamers in the CALU-6 cell line. This would indicate one of three possibilities; 1) the adoption of a slightly different conformation of the pegylated aptamers at 37°C, 2) the structural conformation of the MUC1/Y protein expressed varies slightly in each cell line, or 3) that the PEG offers a varying level of steric hindrance, which may be different depending on the cell line and the other expressed cell surface determinants. The latter two options appear to be more plausible, as if only the thermal stability, hence the conformation of the aptamer, was the determinant of the binding, then similar

cytometric profile changes with the unmodified aptamers would be have been observed, especially in the case of the S51b and S51b-PEG where there is only a 1°C difference in the T_m value.

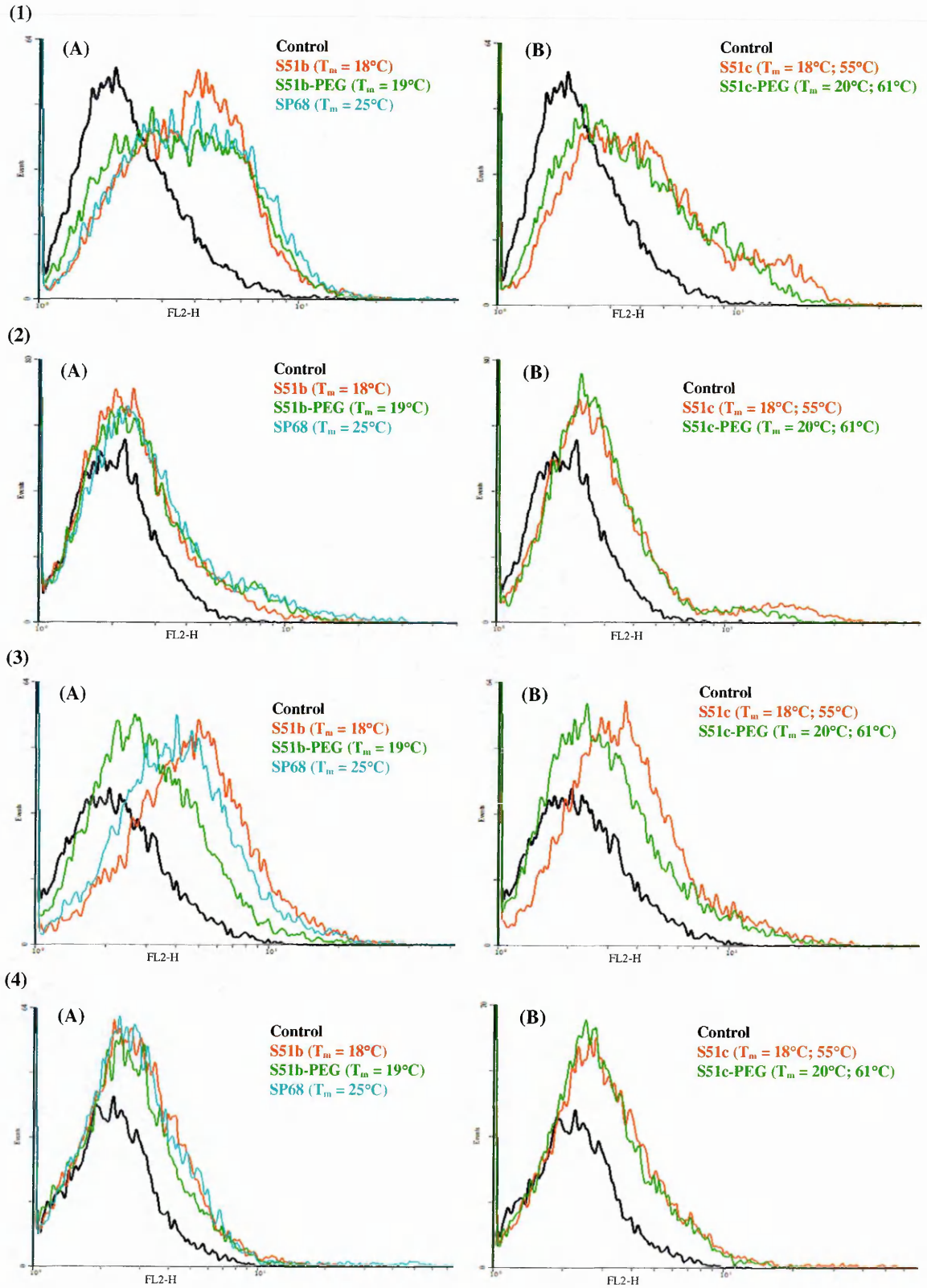


Figure 6.1. FACs histograms of S51 aptamer series and control aptamer SP68. (**A**) S51b version (**B**) S51c version. (**1**) MCF-7 cell line, (**2**) DU145 cell line, (**3**) CALU-6 cell line and (**4**) A498 cell line.

FACs analysis of the S75 aptamer series with four different cell lines is shown in figure 6.2. The S75b aptamer versions all show considerable binding to the MCF-7 cell line, although the S75b-PEG shows fractionally less binding compared to the unmodified aptamer. The S75c aptamer versions appear to display lower binding than the S75b aptamer versions. Again, the S75c-PEG has shown less binding than the unmodified aptamer S75c (figure 6.2, 1A-B). With the DU145 cell line, only limited binding was observed with the S75b and S75b-smd, whilst no binding was detected with the S75b-PEG, S75c and S75c-PEG (figure 6.2, 2A-B). A similar trend that was observed with MCF-7 cell line was also observed with the CALU-6 cell line, where the S75b aptamer versions all displayed significant binding, whilst the S75c exhibits binding but to a lesser extent. Furthermore, the S75c-PEG had an even lower binding than the unmodified aptamer (figure 6.2, 3A-B). No binding of the S75 aptamer series was observed in the MUC1/Y negative A498 cell line (figure 6.2, 4A-B).

The most substantial binding from the FACs results is obtained with the MCF-7 and the CALU-6 cell lines. Furthermore, the S75b-PEG (T_m: 20°C; 40°C) shows equal binding to the unmodified aptamer S75b (T_m: 19°C). This suggests that the adopted conformation as a result of conjugating the aptamer to a large polymer, indicated by the change in T_m value, has not by and large affected the binding potential of the aptamer. However, this does not appear to be so in the case of the S75c-PEG (T_m: 21°C; 35°C; 63°C) where a noticeable reduction in the binding is observed compared to the unmodified S75c (T_m: 22°C; 43°C; 63°C). Hence, a change in the aptamer structure which has led to the change in the second T_m value of the aptamer, from 43°C to 35°C, has resulted in the reduction of the aptamer binding.

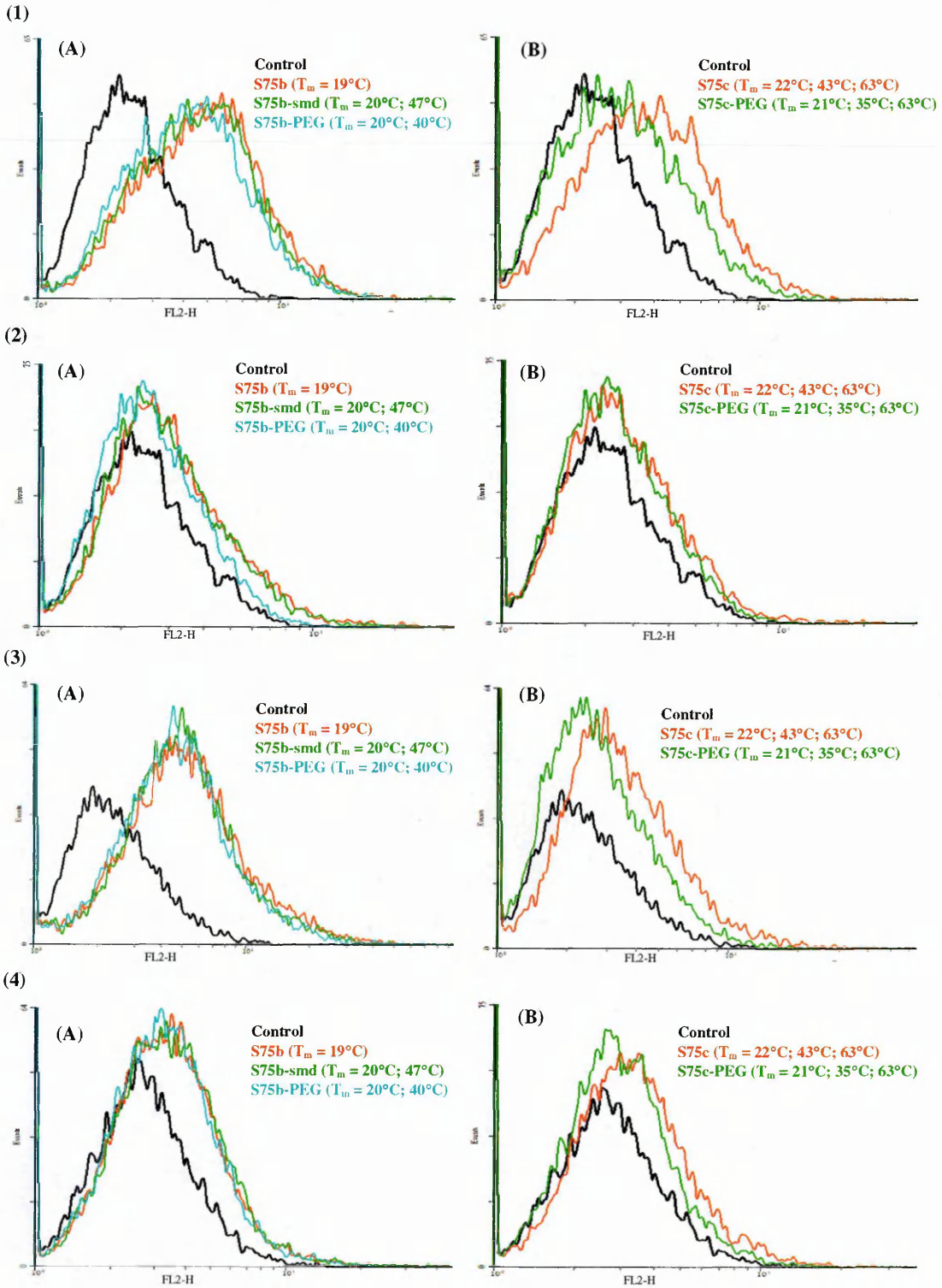


Figure 6.2. FACs histograms of S75 aptamer series. (**A**) S75b version (**B**) S75c version. (**1**) MCF-7 cell line, (**2**) DU145 cell line, (**3**) CALU-6 cell line and (**4**) A498 cell line.

FACs analysis of all three aptamer series has revealed their specificity for the MUC1/Y protein expressed on the MUC1/Y positive MCF-7, DU145 and CALU-6 cell lines, as evidenced by the lack of binding to the MUC1/Y negative A498 cell line. Varying degrees of binding affinity were observed for

each aptamer in the different cell lines, whereby MCF-7 and CALU-6 showed the greatest binding, whilst the DU145 showed the least. The substantially reduced binding observed in the DU145 cells may indicate lower expression levels of the MUC1/Y protein on this cell line. Moreover, with respect to the MCF-7 and CALU-6 cell lines, the 'c' versions (loop region) of S51 and S75 aptamer series showed lower binding than the 'b' version (central variable region). This is likely to indicate that the open structure (low duplex formation) of the 'b' version aptamers as opposed to the loop region (higher duplex content) is preferred for binding to the intact protein on the cells as was also demonstrated by the DNA thermal denaturation studies with the short peptides. Additionally another point to take into consideration is that the T_m values obtained were of the aptamers which were unmodified or modified with PEG. However, the aptamers analysed by FACs have an additional modification, a Cy3 fluorophore, whose effect, if any, on the aptamer conformation is unknown.

6.2 APTAMER CELLULAR INTERNALISATION WITH FLUORESCENCE MICROSCOPY

Each Cy3 fluorescently labelled aptamer of the S11 series (S11b, S11b-smd, S11b-PEG, S11c, S11c-smd, and S11c-PEG), the S51 series (S51b, S51b-PEG, S51c, and S51c-PEG), the S75 series (S75b, S75b-smd, S75b-PEG, S75c and S75c-PEG) and the control aptamers SP68 (selected against a non-MUC1 tumour target peptide) and S2.2 (selected against MUC1) were incubated with MCF-7, DU145, CALU-6 and A498 cell lines in serum-free cell media for 15 minutes, 1 hour and 3 hours at 37°C. Following incubation, the cells were washed and fixed with paraformaldehyde and the nucleus of the cells was stained with Hoeschst 33258. Cellular internalisation of the S11b in MUC1/Y positive MCF-7 cells at incubation times 15 minutes, 1 hour and 3 hours is shown in figure 6.3, where the control is the cells only and the nucleus is stained in blue. Incubation of the cells with S11b at 15 min is comparable to the control, and cellular binding has not taken place during this short incubation period. However, incubation at 1 hr has shown aptamer binding mainly on the cell surface, indicated by a speckled red pattern on the outer edge of the nucleus. Moreover, incubation with S11b for 3 hr has allowed aptamer internalisation to take place and, as seen in many cells, the aptamer appears to have entered the nucleus (indicated by the yellow arrow).

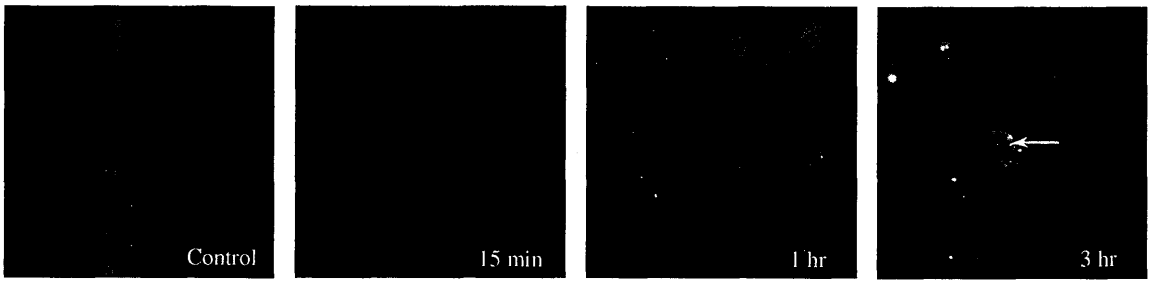


Figure 6.3. Cellular internalisation of S11b in MCF-7cells. MCF-7 cells treated with Cy3-labelled S11b at 15 minutes, 1 hr and 3 hr in serum free media at 37°C.

As all the aptamers exhibited similar fluorescence patterns at 15 minutes and 1 hour in each cell line, for the simplicity of comparison and discussion, only images of the 3 hour aptamer incubation will be presented, analysed and discussed. The fluorescence binding and internalisation of the S11b/c, S51b/c and S75b/c aptamer series and the control aptamers in MCF-7 cells is shown in figure 6.4. A high uptake of the aptamers S11b, S51c, S75b, S75c and SP68 was observed, whilst very little uptake was observed with the 3 scrambled aptamers (S11b-smd, S11c-smd and S75b-smd) and the MUC1 aptamer. However, with the exception of the S11c-PEG, the other pegylated aptamers showed extensive uptake compared to their unmodified sequences.

The aptamers studied for fluorescence binding and cellular internalisation in MCF-7 cell line were also studied using the DU145 cell line (figure 6.5). A relatively high uptake of S11c, S51c, S75b and S75c, mainly the loop region aptamers, was observed, whilst the central variable region aptamers S11b and S51b, the 3 scrambled aptamers, and the control aptamers SP68 and MUC1, all show considerably low aptamer binding and uptake. S11b-PEG, S51b-PEG, S51c-PEG and S75b-PEG show binding and internalisation that is comparable to that of the unmodified aptamers and not as extensive as with the MCF-7 cell line. S11c-PEG and S75c-PEG shows lower uptake than the unmodified aptamers S11c and S75c.

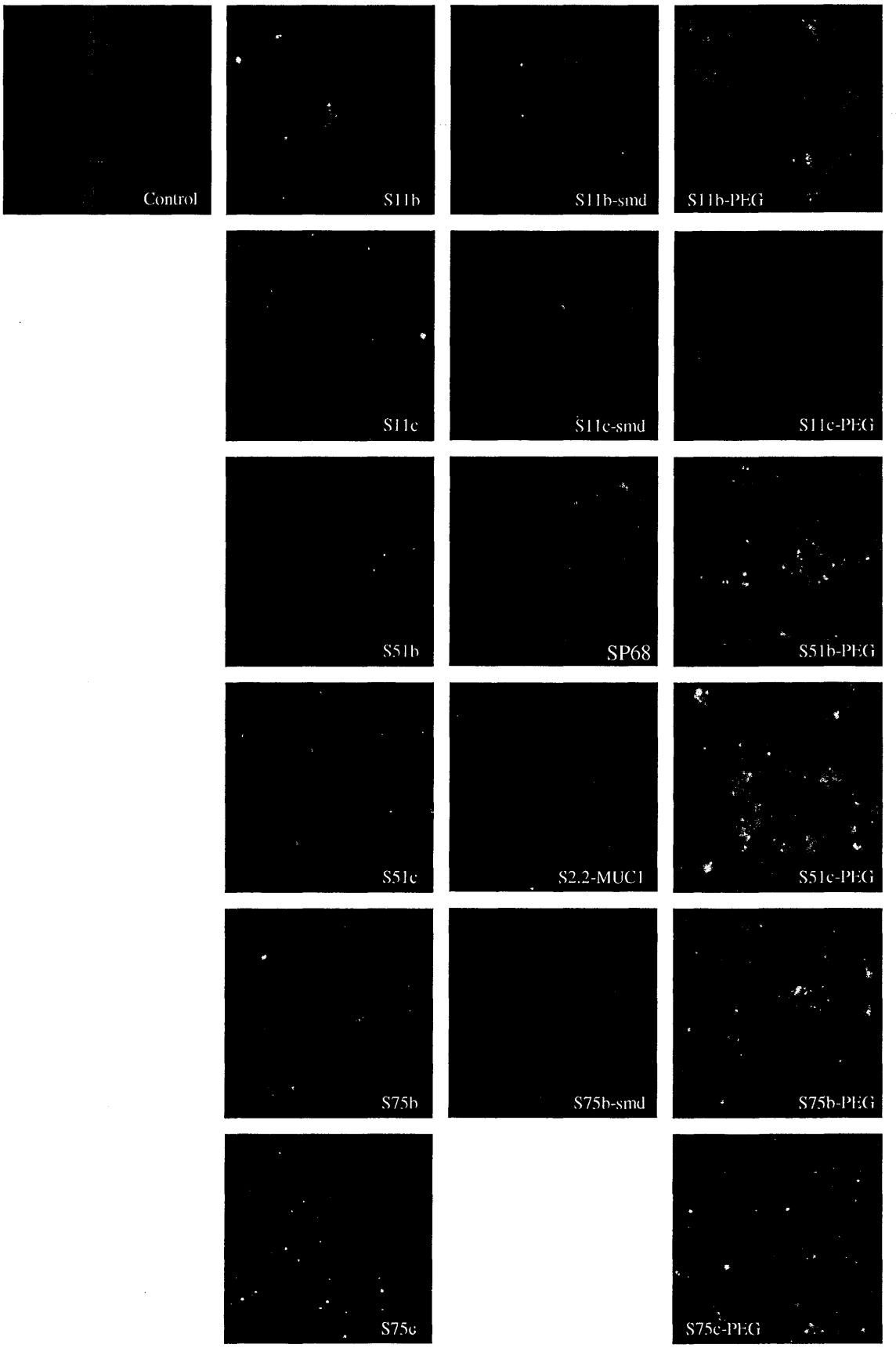


Figure 6.4. Cellular internalisation in MCF-7 cells. MCF-7 cells treated with Cy3-labelled MUC1/Y aptamer, scrambled aptamers and control aptamers in serum free media at 37°C for 3 hr.

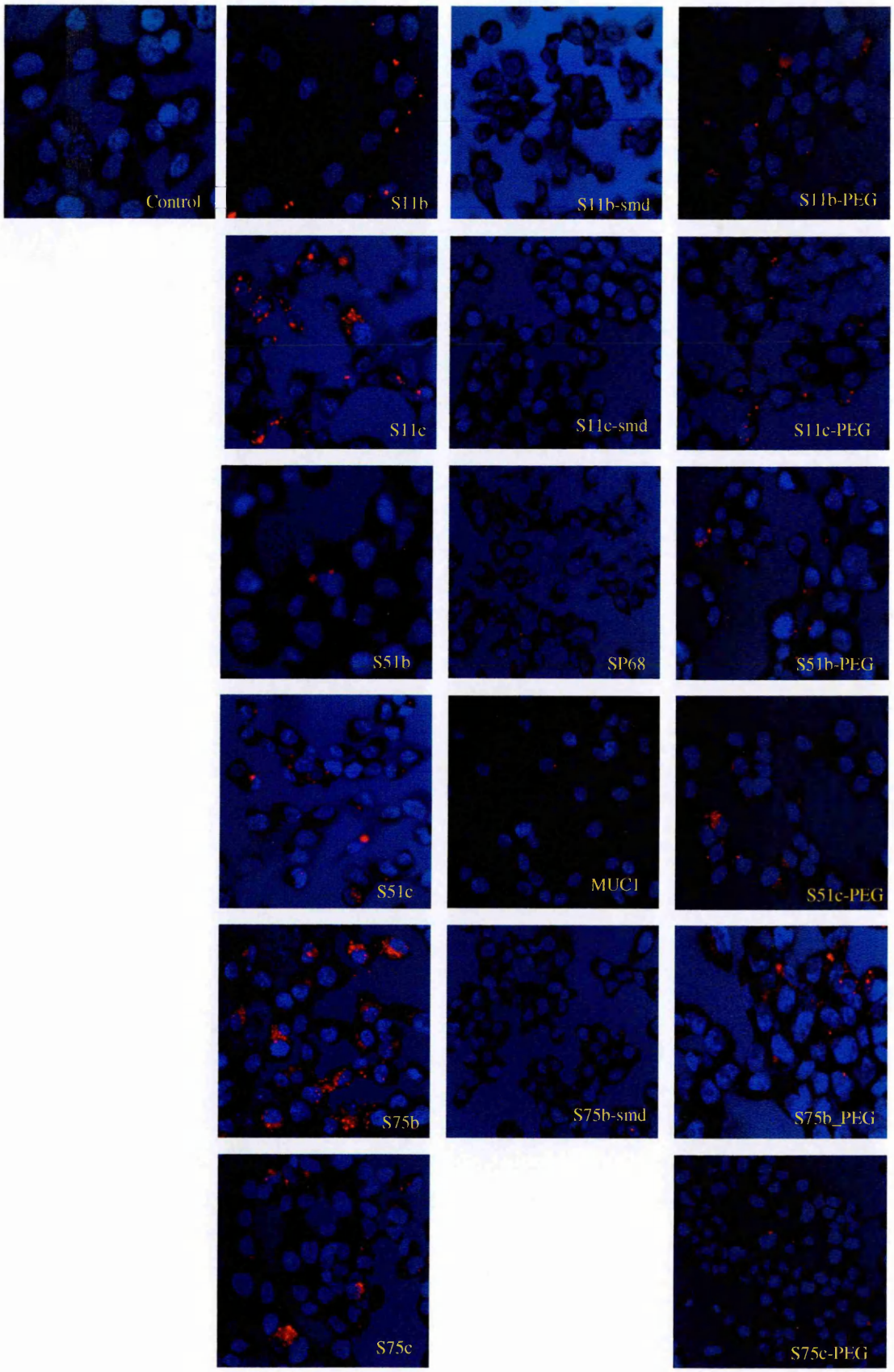


Figure 6.5. Cellular internalisation in DU145 cells. DU145 cells treated with Cy3-labelled MUC1/Y aptamer, scrambled aptamers and control aptamers in serum free media at 37°C for 3 hr.

The aptamers studied using the CALU-6 cell line are shown in figure 6.6. Aptamers S11b/c, S51b, S75b/c, S11c-smd, S75b-smd and SP68 all show significantly high uptake into the cells whilst very little uptake was observed with the S11b-smd, S51c and MUC1 aptamers. With the exception of the S51b-PEG, the other five pegylated aptamers all exhibited very high uptake in the cells. The uptake of the aptamers in the CALU-6 cell line has demonstrated an atypical pattern to the norm that is usually observed in a speckled manner within the cytoplasm. With this particular cell line the uptake of the aptamer appears to be more in an aggregated form.

The fluorescent binding and cellular internalisation of the aptamers using the MUC1/Y negative A498 cell line is shown in figure 6.7. The MUC1/Y aptamers, including the pegylated and scrambled aptamers all showed no significant binding to the A498 cells. However, noticeable uptake of aptamers MUC1 and SP68 was displayed in the cytoplasm.

The fluorescence binding and internalisation studies using fluorescence microscopy has shown that the MUC1/Y aptamers have specificity to their target protein, as this was clearly indicated by the lack of binding observed in the MUC1/Y negative cell line. The low binding and therefore no internalisation of the MUC1 aptamer in the MUC1/Y positive cell line possibly indicates that the expression of the MUC1 protein was not as significant to be detected as was the MUC1/Y protein. The SP68 aptamer displayed substantial binding in the all three MUC1/Y positive cell lines, however, the SP68 also showed binding and internalisation in the MUC1/Y negative A498 cell line. This could suggest that the protein SP68 was selected against, also a known tumour marker, is likely to also be expressed in these cell lines. From the six unmodified MUC1/Y aptamers, S75b and S75c are the two aptamers which have consistently shown high binding and internalisation in all three positive cell lines. Furthermore, typically it has been observed that if binding of the unmodified aptamer is observed, then greater binding and uptake of the pegylated aptamer for the same cell line is obtained. This indicates that the large polymer does not affect the binding of the aptamer to the target protein nor does it cause non –specific binding, but helps facilitate internalisation of the aptamer once bound.

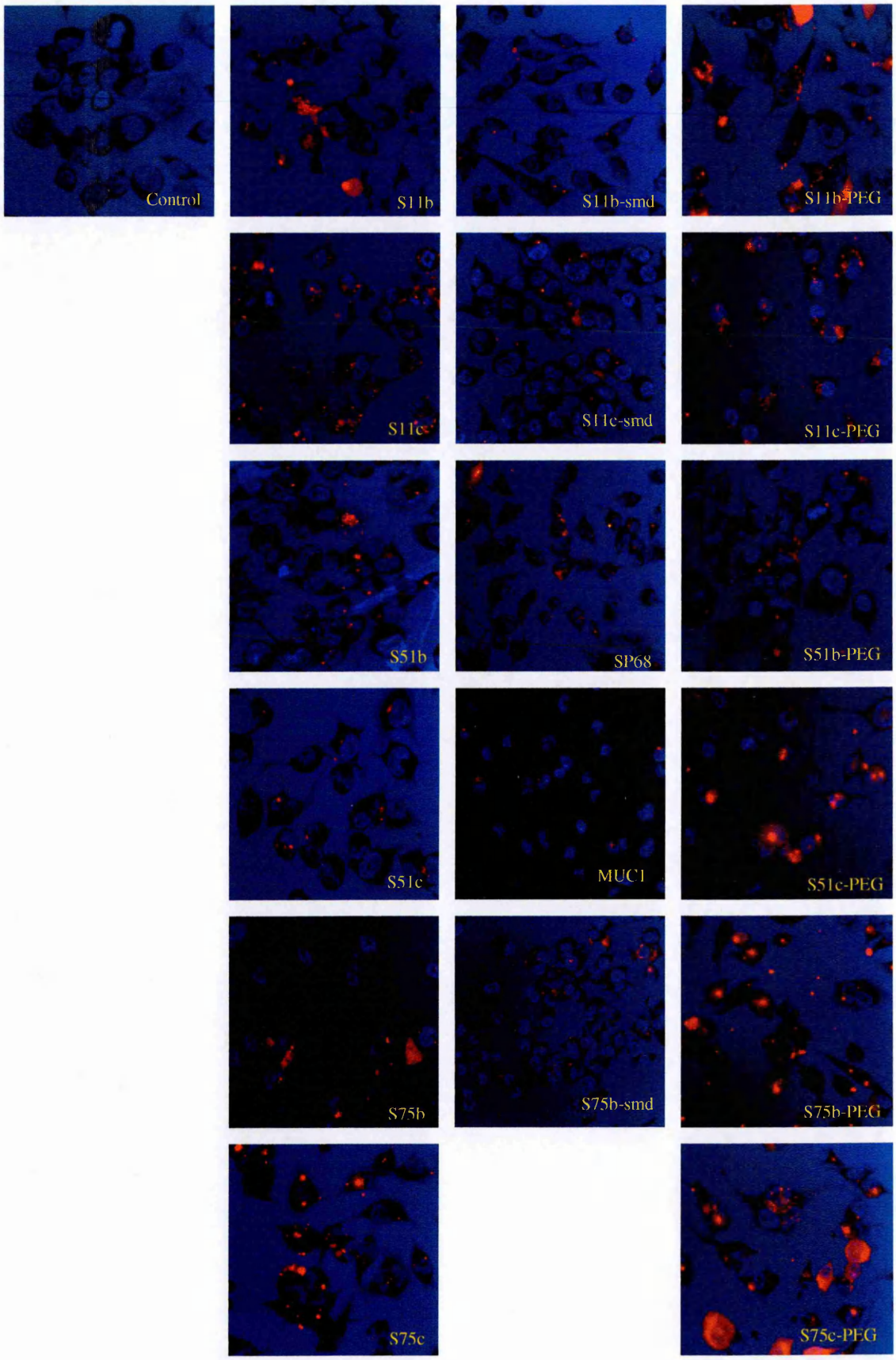


Figure 6.6. Cellular internalisation in CALU-6 cells. CALU-6 cells treated with Cy3-labelled MUC1/Y aptamer, scrambled aptamers and control aptamers in serum free media at 37°C for 3 hr.

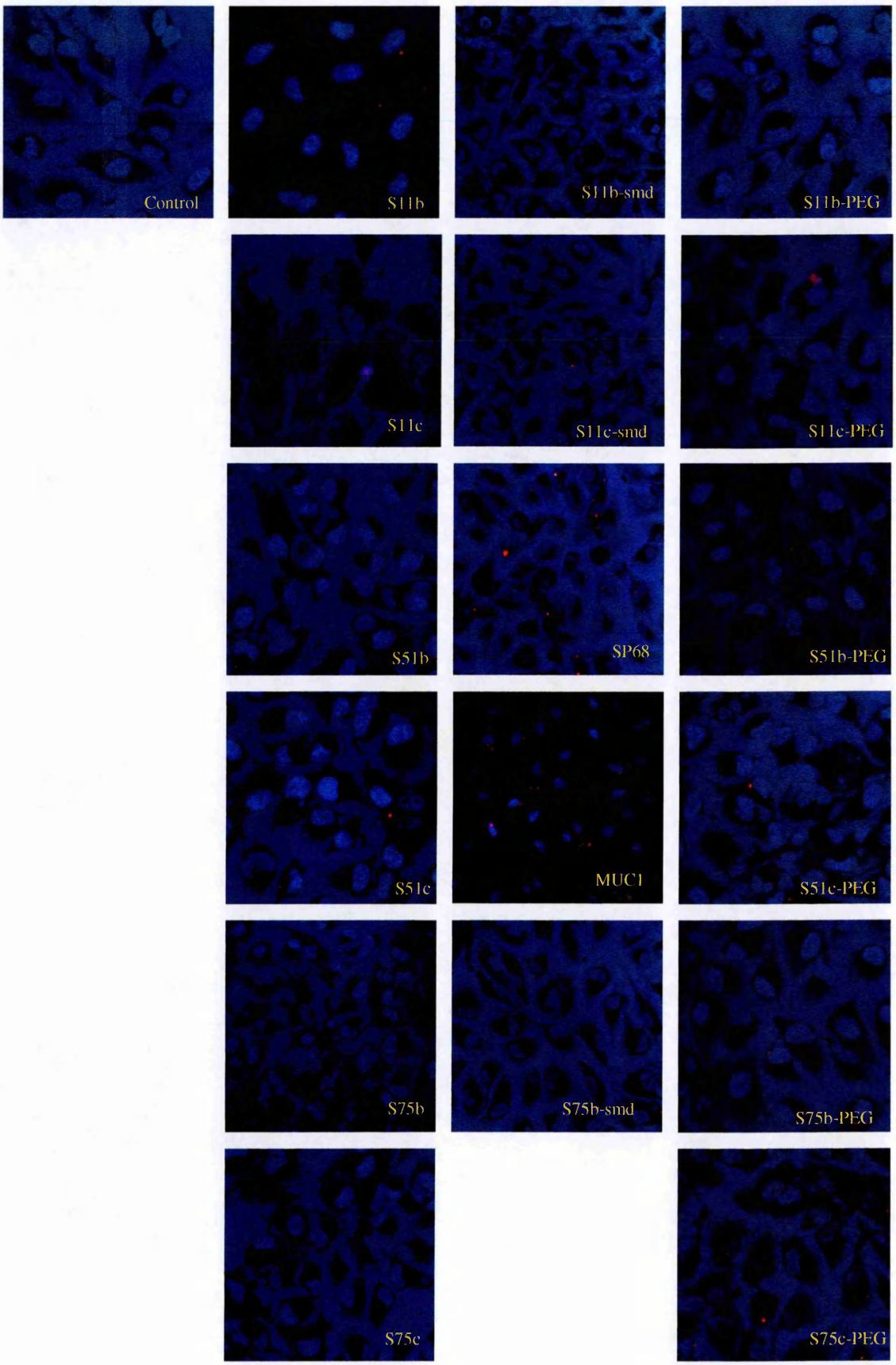


Figure 6.7. Cellular internalisation in A498 cells. A498 cells treated with Cy3-labelled MUC1/Y aptamer, scrambled aptamers and control aptamers in serum free media at 37°C for 3 hr.

6.3 CONCLUSION

The results obtained from the FACs analysis and the fluorescence microscopy studies both draw a parallel with regards to the MUC1/Y aptamers, showing greatest binding to the MUC1/Y protein expressed on the MCF-7 (human breast carcinoma) and the CALU-6 (human lung carcinoma) cell lines. Moreover, both techniques have shown that the expression of MUC1/Y on the DU145 (human prostate carcinoma) cell line is not substantial, thus extensive binding with the aptamers was not observed. As there has been no antibody available against the MUC1/Y protein, either polyclonal or monoclonal, aptamers can also be used to test for the presence of MUC1/Y on the cell surface using FACS, and thus identify MUC1/Y positive cell lines. However, contradictory results were obtained with regards to the pegylated aptamers, where in the FACs analysis they were shown to have reduced binding compared to their respected unmodified aptamers. However, the analysis of the fluorescence confocal microscopy studies shows that PEG enhances the uptake of the aptamers. This could be attributed to the difference in aptamer incubation periods with the cell lines, where for FACs analysis the aptamers were incubated for 1 hour and for confocal microscopy studies aptamers were incubated for 3 hours. Furthermore, it is possible that although PEG has a slight inhibitory effect in initial binding, possibly due to steric hindrance, once bound it facilitates aptamer internalisation into the cell. Another contrast in the findings between the two techniques was with respect to the control scrambled aptamers. FACs analysis indicates that the scrambled aptamers share comparable affinity for the protein as the non-scrambled aptamers, whereas the confocal microscopy studies show very little uptake of the scrambled aptamers even with the 3 hour of incubation. This discrepancy could be attributed to the different condition used in both techniques during the incubation period. The aptamers were incubated with the cells in PBS for the FACs analysis whilst for the fluorescence microscopy experiments the aptamers were incubated with the cells in serum free media. In any case, this differentiation is desirable, as demonstrates greater specificity of the selected aptamers over their scrambled counterparts.

The fluorescence microscopy provides confidence that the MUC1/Y aptamers have specificity and affinity for their target protein and, in addition, the pegylation of the aptamers enhances their binding and

internalisation further. Whilst all six aptamers have shown binding affinity and specificity, the S75b and S75c aptamers appear to be more promising as their binding has remained consistent in the three MUC1/Y positive cell lines.

CHAPTER SEVEN

FINAL CONCLUSIONS

7.0 CONCLUSIONS

There are many proteins, growth factors, hormones, antibodies and small molecules which are involved either at different stages of cancer development or during the progression of the malignancy. However, involvement of any of these molecules in cell proliferation, initiation of excessive cellular division or cell signalling, indicate their potential as tumour biomarkers. Molecules involved in cell signalling are vital as, through mutations, the normal function is bypassed, allowing cell proliferation to occur infinitely and inhibiting apoptosis, which can result in greater expression of particular proteins. Hence, over-expression of proteins predominantly found in malignant tissues, as opposed to the expression of the proteins present in normal tissues of the same organ, are crucially important and of great interest as potential molecular tumour targeting biomarkers. MUC1/Y is such a protein. The MUC1/Y transmembrane protein, a splice variant of the MUC1 glycoprotein, has been found to be prevalent in many epithelial carcinomas and, unlike MUC1, this protein does not possess the variable tandem repeat region. This has resulted in the absence of the carbohydrate structures which are characteristically present in MUC1, causing substantial deficiency in the polarisation of the protein, permitting the MUC1/Y to be greatly overexpressed in the cell surface. Furthermore, the loss of the adhesive properties from the absence of the carbohydrate structures has also contributed to the aggressive nature of the tumours expressing this protein, by enhancing their ability to invade into the lymph vessels and metastasise to the other organs. Therefore, the development of low toxicity molecules with the ability to bind to MUC1/Y with substantial affinity and specificity, which could potentially result in tumour cell death through interfering with the cell signalling pathways, is a viable approach for target cancer therapy.

Following this rationale, MUC1/Y was selected as a potential therapeutic target and as a tumour biomarker for the diagnosis of many epithelial tumours. As MUC1/Y is a transmembrane protein, and due to the complexities involved in its isolation and purification, the whole protein was not employed for the selection of the aptamers. Instead two short synthetic peptides were designed to be used as targets for the selection of MUC1/Y aptamers. The MUC1/Y peptides consist of 10 and 20 amino acids, referred to as 10mer peptide and 20mer peptide respectively. The 10mer peptide sequence (TEKNAFNSSL) was selected on the basis of 5 amino acids before the splice site of the MUC1 and 5 amino acids after the

splice site of the protein which generates the MUC1/Y from the MUC1 extracellular domain protein sequence. The 20mer peptide sequence (SVPSS**TEKNAFNSSL**EDPST) is an extension to the 10mer peptide of 5 additional amino acids either side. The purpose of the MUC1/Y peptides was to mimic the splice site present in the MUC1/Y protein, which is a specific portion of the amino acid sequence that differentiates the MUC1/Y isoform from the MUC1 protein and its other splice variants.

Taking into consideration the traditional SELEX system, involving numerous rounds of selection and amplification, an alternative approach for generating aptamers to potentially any target was considered. SimpLex, a novel one round selection method developed to obtain aptamers against the MUC1/Y peptides entailed immobilisation of the biotinylated peptides to the streptavidin coated PCR tubes to which the single stranded combinatorial library was incubated with the target peptides. The bound aptamers were separated according to their binding strength to the target peptides via either a salt gradient or a heat gradient elution. Forty sequences were obtained from both methods and peptides, of which only three aptamer sequences (S11a, S51a, S75a) were considered for further investigation. These three sequences were chosen to be taken further on the basis that they were also obtained from a repeat selection employing the traditional affinity chromatography system. Additionally, another six sequences were designed from the parent sequences which consisted of two different constructs. The first construct involved the removal of the hybridising primer sequences leaving only the central variable region (S11b, S51b, and S75b) and the second construct entailed the sequence forming the stem-loop conformation obtained from the predicted secondary structures (S11c, S51c, and S75c). SimpLex has shown potential to generate aptamers by acquiring binding oligonucleotides of the same sequences as those obtained from the traditional selection system. Hence, SimpLex presents a possible novel alternative to the SELEX process, eliminating numerous successive rounds of selection and PCR amplification cycles over a period of several days, with the additional advantage of using low amounts of target and combinatorial library. Furthermore, when compared to other single step selection methods such as the MonoLex and ASExp of which both aim to remove the repetition of successive rounds of selection, both processes appear to be far more time consuming than the SimpLex method. The MonoLex system involves incubating the combinatorial library with the target attached in an affinity capillary column. The unbound

oligonucleotides are washed away with numerous buffer washes, followed by physical segmentation of the column, where the bound sequences are removed via desorption for PCR amplification and sequencing. In this system the opportunity to obtain sequence homology is quite limited [199]. The ASExp system entailed PCR amplification of the ssDNA library to generate a dsDNA library, which was incubated with the target causing denaturation of the dsDNA upon the ssDNA-target complex formation. Separation of the bound sequences from the unbound was performed using magnetic beads attached with random ssDNA, resulting in a complex of the bound sequence-target-magnetic bead, which was PCR amplified for sequencing. The entire process approximately takes ten hours to complete [206]. Thus, the SimpLex offers a practical yet much simpler, fast and less laborious process for isolating the greater binding sequences from the low binders.

The unmodified aptamers are subjected to rapid renal clearance which prevents the aptamers from reaching the biomarker and achieving an adequate therapeutic response. Thus the aptamers are required to be modified so as to surpass the minimum molecular weight cut off for glomerular filtration of 30 kDa. This limitation was addressed by chemically modifying the potential aptamers through a process of pegylation, which entailed conjugating the truncated aptamers to a 32 kDa molecular weight polymer, polyethylene glycol (PEG). Several different parameters for aptamer conjugation were investigated before the optimised condition was obtained to provide a high yielding aptamer pegylation reaction of 79%. The optimal reaction conditions are; 15: 85 (v/v) mixture of 53.4 mM sodium carbonate buffer (pH 10) and DMF, at 60C for 2 hours. These pegylation conditions significantly varied from those described in literature where Kang et al [226] utilised pegylation conditions of 100 mM sodium carbonate buffer (pH 8.5), 5 molar excess of PEG, in equal volume of acetonitrile with a reaction time of 1 hour. This pegylation method has also been described by Healy et al [221] with the difference of 2.5 molar excess of PEG. Sainathan et al [238] employed different conditions to some extent for pegylation which included PBS (pH 8.2), 5 molar excess of PEG, at 25°C for 30 minutes. These described methods, however did not indicate the percentage yield of the pegylated product obtained and were unsuccessful in achieving significant pegylation of the MUC1/Y DNA sequences. The presence of pegylated aptamers was confirmed by anion exchange HPLC analysis, native and agarose gel electrophoresis.

The pegylated sequences were characterised alongside the unmodified sequences for any changes in the affinity or specificity of the potential aptamers for their target peptides upon conjugation. With the aim of the selected aptamer to be investigated further as therapeutic agents, it was vital that the potential aptamer was pegylated. Therefore, to avoid inaccurate conclusions with respect to which aptamer may have the greatest potential for therapeutic development, the selected oligonucleotides were screened using qualitative techniques in parallel with the pegylated potential aptamers. This allowed the concluding of which aptamer retained its activity subsequent to pegylation and which would be the most suited candidate to be taken for further studies.

The qualitative techniques employed to assess the individual aptamer affinity and specificity for the MUC1/Y 10 mer and 20 mer peptides were EMSA, affinity chromatography and DNA thermal denaturation. The full length, truncated and pegylated aptamers analysed by EMSA all exhibited marginal band shifts in the native PAGE when incubated with the 10 mer and 20 mer peptides as compared to the band of the free aptamer. This was due to the MUC1/Y peptides having a low molecular weight (1 – 2 kDa), and not significantly altering the electrophoretic properties of the aptamer to result in a retardation within the gel and produce a visible band shift. With respect to many of the sequences, such as the S51a, a marginal retardation was observed with the S51a-20 mer mixture whilst the S51a-10 mer mixture was running slightly faster than the free S51a oligonucleotide. On the other hand, the reverse was observed with the S11a, where S11a-10 mer mixture migrated slightly ahead and the S11a-20 mer mixture was fractionally retarded than the free S11a oligonucleotide. Typically, it is expected that upon aptamer-peptide complex formation a slower mobility of the complex is observed compared to that of the free aptamer. However, this is not always the case, as upon the aptamer binding to the target, the change in conformation of the aptamer can affect its migration through the gel. For example, if an aptamer is induced to have a greater secondary structure upon binding to the peptide, then this may lead to the aptamer-peptide complex having a swifter mobility than the free aptamer, whilst if an aptamer is induced to adopt an open structure upon binding, then the aptamer-peptide complex may be slightly retarded in the gel. Furthermore, the heating of the aptamers in the presence of the MUC1/Y peptides to 95°C for 5 minutes appeared to interfere with the reformation of the aptamers secondary structure. This effect was

indicated in the gel by the disappearance of the aptamer bands after staining with EtBr, an intercalating dye, and this was observed for all the aptamers analysed including the control aptamers. Thus, suggesting that non-specific aptamer binding to the target when fully denatured (presence of no secondary structure) is likely to take place, particularly as the control aptamers are of the same base composition as the selected ones. Therefore, due to this lack of sensitivity, the EMSA technique cannot to be used to deduce which aptamer retains to a greater extend its secondary conformation. In addition, this technique cannot be used to conclude with certainty that the marginal band shifts observed were a result of the aptamer-peptide complex formation, as this can also be attributed to the possible discrepancies of the technique. Nonetheless, upon heating of the aptamers, their stabilisation in the unstructured conformation by the peptides was also observed in the thermal denaturation studies.

DNA thermal denaturation studies of the aptamers have shown substantial decreases in fluorescence during the second and third thermal denaturation cycles of the aptamer in the presence of the MUC1/Y peptides. This effect indicated that binding to the peptides had occurred and the peptides possessed the ability to inhibit the renaturation of the aptamers, causing the potential aptamers to remain unstructured. It was also observed that the 20 mer peptide had a greater effect on stabilising the unstructured aptamer conformation than the 10 mer peptide. However, as the aim is to utilise the aptamer as a therapeutic agent, it was more relevant to consider the binding of the MUC1/Y peptides with the aptamer prior to any heating, as the highest temperature the aptamer would be subjected to in *in vivo* would be the physiological temperature of 37°C. For all three series (S11, S51 and S75) it was observed that truncation of the full length aptamers to the central variable region (b series) resulted in significant decrease in the thermostability, whilst truncation to the loop region (c series) resulted in either change in conformation (S11c) or slightly enhanced thermostability (S51c and S75c). The pegylation of the b series aptamers has resulted in a change in conformation and enhanced thermostability (S11b-PEG and S75b-PEG) with the exception of S51b-PEG, which showed no notable difference. The pegylation of the c series oligonucleotide has resulted in a slight increase in the thermostability of S11c-PEG and S51c-PEG, whilst the S75c-PEG showed a slight decrease. The S11b-smd exhibited a similar melting profile and temperature as the S11b (21°C), which could be explained by the predicted m-fold secondary structure

showing identical number of GC (2) and AT (1) base pairs in the single stem conformation of both oligonucleotide. However, the S11c-smd (47°C) and S75b-smd (20°C; 47°C) showed a change in the melting profile and temperature compared to S11c (21°C; 45°C) and S75b (19°C) respectively. According to the predicted structure for S11c, the aptamer possess two stems separated by a small bulge between the stems containing 3 GC and 4 AT base pairs, thus, possibly resulting in two different T_m values. Whereas the S11b-smd has a single straight stem conformation consisting of 4 GC and 1 AT base pairs and as the stem is largely composed of GC base pairs, this is perhaps the reason for a single T_m value being observed. The S75b has a short single stem comprised of 1 GC and 2 AT base pairs providing a single T_m value, whilst the S75b-smd forms a stem of 3 GC and 2 AT base pairs resulting in two different T_m values. The varying thermostability of the free aptamers observed from the different modifications performed suggests that each aptamer possesses its own individual characteristics. However, upon the addition of the 20 mer peptide, with the exception of S51c-PEG, all potential aptamers have displayed a T_m value of below 37°C. Conversely, this is not the case with the 10 mer peptide, where T_m values above 37°C are detected along with greater stabilisation of many of the aptamers. This shows that the length of the peptide makes a remarkable difference in the conformation the aptamer adopts and its thermostability. This investigation has shown that a secondary structure is preferred for the initial binding. However, upon binding, if the aptamer adopts a more labile conformation, this further increases the affinity of the aptamer for its target, presumably as a change in structure occurs upon binding. If an aptamer lacks sufficient secondary structure, it was observed that specificity is lost. This was also displayed by the analysis of the aptamers through affinity chromatography.

The affinity column chromatography technique results showed no elution of the bound aptamers even at salt concentrations as high as 1.6 M NaCl, even though different amount of aptamers were retained by the column. This was characteristic of a tight binding of the aptamer, though the lack of specificity may be due to the low melting temperatures of the aptamers. Thus, it is likely that during this study, the aptamers were largely in an unstructured conformation at ambient temperature, which resulted in a loss of specificity, indicated by the binding of the control aptamers to the MUC1/Y peptides to some extent.

However, in accordance with the thermal denaturation studies, the pegylated oligonucleotides displayed increased binding for both the 10 mer and 20 mer peptides, which further suggests that an initial secondary structure conformation of the aptamer is vital for increased affinity. A point to note with respect to the results obtained from the affinity column chromatography technique is that they may vary from the results obtained via the EMSA and the thermal denaturation studies due to the difference in the incubation condition. The EMSA and the thermal denaturation techniques involved incubation of the potential aptamer with the peptides free in solution, whereas the affinity column chromatography method required the potential aptamers to be incubated with the peptides in a stationary phase. The qualitative characterisation methods described here are not conventional methods which can be found typically in literature to assess the binding of the aptamers to the selected target, thus comparison to previous work cannot be made. These characterisation techniques are subjected to providing inaccurate results due to the inconsistencies associated with the nature of the techniques. To accurately determine the biomolecular interactions of the aptamer with the target and obtain real-time kinetic analysis such as rate constants and binding affinities, surface plasmon resonance (SPR) spectroscopy should be employed [239]. SPR further allows the opportunity to explore the effects of different salt concentrations on the binding along with the possibility to gain a better understanding of the aptamer-target complex formation and the stability of the complex via competition and displacement experiments [240].

Although the qualitative characterisation techniques show variability in the affinity and specificity of the potential aptamers for their target peptides, which was much dependent upon the conditions employed in the individual studies, the subsequent cell binding and internalisation studies have shown greater clarity of the prospective aptamers with respect to their specificity and affinity for the full, intact, MUC1/Y protein expressed on the cell surface. Cell binding studies using FACs analysis has shown significant binding of all the potential aptamers to the MUC1/Y positive MCF-7 and CALU-6 cell lines, whilst reduced binding to the MUC1/Y positive DU145 cell line. No binding of the potential aptamers was observed in the MUC1/Y negative A498 cell line. The S11b-smd (T_m: 21°C) has the same T_m value as S11b (T_m: 21°C) but showed a marginal reduction in the binding at the FACS analysis, whilst S11c-smd (T_m: 47°C) has a different conformation to the S11c (T_m: 21°C; 45°C) and also exhibited a small decrease

in binding with the MCF-7 cell line. The S11b-PEG displayed similar binding to the unmodified S11b whilst the S11c-PEG exhibited a decrease in binding compared to S11c with the MCF-7 cell line. However, the S11 series sequences displayed comparable binding to each other within the series with the DU145 and CALU-6 cell lines. The S51 series oligonucleotides demonstrates similar binding to each other in the MCF-7 and DU145 cell lines, which could be attributed to their comparable T_m values. However, this was not the case with the CALU-6 cell line where S51b-PEG (T_m: 19°C) and S51c-PEG (T_m: 20°C; 61°C) have shown reduced binding compared to the unmodified oligonucleotide S51b (T_m: 18°C) and S51c (T_m: 20°C; 61°C). Hence, the variation in binding observed for the same oligonucleotides in different MUC1/Y positive cell lines further indicates that the actual expression of the MUC1/Y protein also dictates the binding potential of the prospective aptamers. In addition, the SP68 oligonucleotide has shown comparable binding to the S51 series aptamers, although this is expected, as the SP68 was selected against a different tumour marker, which is likely to be expressed on the surface of these malignant cells. The series S75b oligonucleotides show comparable binding to each other in all three MUC1/Y positive cell lines even though the S75b-smd (T_m: 20°C; 47°C) and the S75b-PEG (T_m: 20°C; 40°C) have different T_m values and conformations to the S75b aptamer (T_m: 19°C). In contrast, the S75c-PEG (T_m: 21°C; 35°C; 63°C) has shown slightly lower binding in the MCF-7 and CALU-6 cell lines compared to the unmodified S75b oligonucleotide (T_m: 22°C; 43°C; 63°C). The positive shift in the histograms of the S11b and S75b oligonucleotides are comparable to those observed by the anti-MUC1 aptamer in an individual study [235] with the MCF-7 cell line. Hence, supporting the hypothesis of the oligonucleotides generated against the MUC1/Y peptides may be capable of recognising the peptide epitope which would be present within the full protein expressed on the cell surface.

A similar trend to that observed in the FACs analysis is also observed with the fluorescence microscopy studies, where greater binding and internalisation of the potential aptamers was detected with the MCF-7 and CALU-6 cell lines than the DU145 cell line, whilst no binding was obtained with the A498 cell line. Nonetheless, substantial differences were observed with individual prospective aptamers, as the scrambled oligonucleotides; SP68 and MUC1 showed very little binding, if any, with the three MUC1/Y

positive cell lines. The pegylated oligonucleotides displayed far greater binding and internalisation than the unmodified oligonucleotides, without resulting in any non-specific binding. The fluorescent signal for cellular binding and internalisation of the potential MUC1/Y oligonucleotides with the MCF-7 cell line is shown to be similar to the MUC1 aptamer, whilst the pegylated sequences have presented fluorescent signals that are alike, produced by the multi-walled carbon nanotubes (MWNT) – MUC1 aptamer complex. However, in the same study, the MUC1 aptamer has shown very little uptake by the CALU-6 cell indicating the cell line to be MUC1 negative [241]. Conversely, this was not the case for the potential MUC1/Y aptamers, as considerable cellular binding and internalisation was observed suggesting that the CALU-6 cell line is MUC1 positive. This discrepancy in cellular binding and internalisation of the same cell line but of a different batch is observed often, thus inconsistency in the protein expression levels can be expected and should be not presumed that aptamer binding has not taken place.

The differences in binding outcomes presented by the two techniques, FACs and fluorescence microscopy can be attributed to the different conditions utilised during the incubation of the aptamers with the cells. In the FACs technique the aptamers were incubated with the cells in PBS, whilst for the fluorescence microscopy the aptamers were incubated in serum free media. Thus, it is evident from the different biophysical and *in vitro* cell binding studies performed that the affinity and specificity of the each oligonucleotide varies depending upon the conditions employed in the implementation of the technique. On this basis the results obtained from the fluorescence microscopy studies can be considered to be the most realistic, as the serum free media used in this technique is the closest in mimicking the physiological conditions. Therefore, as the pegylated oligonucleotide did not show any reduced binding or adverse effect on the specificity of the aptamers towards the MUC1/Y protein, but rather enhanced the existing properties of the aptamers, it would appear feasible to take the pegylated oligonucleotide further for characterisation studies to determine which of the six pegylated oligonucleotides would be the best candidates for clinical development.

Finally a point to note, although the experimental results from the FACs analysis and fluorescence microscopy studies are encouraging, the varied and less definitive outcomes obtained from the qualitative

characterisation studies cannot be ignored. The disparity between the results obtained the in vitro internalisation and qualitative studies could indicate a potential failure within the selection method. This is further indicated by the lack of sequence consensus obtained from the SimpLex selection method. The following could propose to be significant validations for the inefficient outcome obtained from the selection method; physiological conditions (buffer and temperature), negative and counter selection were not incorporated during the SimpLex selection. In addition, SimpLex only consisted of sequencing 10 clones from each of the two methods (salt and temperature gradient) and the two target peptides (10 mer and 20 mer), where this low number of sequences cloned is not sufficient to increase the possibility of achieving sequence homology or to validate a new selection method. For rectification of the SimpLex protocol to achieve an improved and more accurate end result for future targets, the following should be taken into consideration; cell media should be employed as the physiological solution in which the oligonucleotide library is incubated with the target, at a physiological temperature of 37°C. The cell media solution can be considered to be comparable to the physiological fluid, as this media is also utilised to grow cells. This change alone should aid in eliminating unstructured low binding sequences at 37°C and enrich the pool of oligonucleotide sequences with a stable conformation under correct buffer conditions. Negative and counter selections should be performed, for example, the oligonucleotide library should be subjected to the support medium and to molecules similar to that of the target respectively, in order to remove non-specific sequences and increase the stringency of the selection. In the case of aptamer selection against the MUC1/Y targets, the library should be incubated with the streptavidin coated tubes to performing a negative selection and to the MUC1 peptides for perform counter selection. Furthermore, to increase the probability of obtaining sequence homology, typically 100 individual clones should be sequenced for each method and peptide [242,243].

7.1 FUTURE WORK

The qualitative characterisation and the *in vitro* cell binding studies have given a good insight as to how the binding potential of the aptamer is altered depending upon the conditions utilised in the investigation. The results obtained in this investigation are preliminary, but leave great scope for aptamer development

as a potential cancer therapeutic agent. This work also provides a good foundation for future work which would entail performing quantitative biophysical binding studies using conditions which are closely related to the physiological. This could include using cell media as the solution in which the aptamer and peptide are incubated in, rather than the use of binding solution or buffers. In addition, all the studies should be performed in a controlled environment with a temperature of 37°C, as it was obvious from the current studies that a lot of the aptamers presented melting temperatures below this value and would be potentially unstructured at physiological temperatures. Incorporation of these conditions during aptamer characterisation studies will ensure the natural conformation of the aptamer at 37°C is adopted and hence reduce the potential variability that was observed in these studies in the aptamers' affinity and specificity for their target. Furthermore, with respect to the biophysical characterisation studies in particular, the thermal denaturation experiment where a 1:1000 aptamer: peptide ratio was used and could be repeated using lower ratios of peptide to aptamer, to determine the minimum concentration of peptide required to inhibit the renaturation of the aptamer and reduce non-specific binding.

Following aptamer characterisation studies cell cytotoxicity of the pegylated aptamers would ascertain the aptamers' potential as a therapeutic agent and provide an indication of which aptamers are able to result in cell apoptosis. Results from this study would also aid in narrowing down the potential therapeutic candidates from six to two.

Upon selecting the pegylated aptamer to be taken forward for further investigation, it would be informative to deduce the structure of the aptamer free in solution and the conformation it adopts upon binding to the peptide. Structural analysis of the aptamer can be performed using nuclear magnetic resonance (NMR) and x-ray crystallography techniques. The NMR studies would aid in the resolving the three-dimensional structure of the peptide, aptamer and the peptide-aptamer complex at atomic resolution. As the NMR structural studies are performed in solution, the experimental conditions can be generated to mimic physiological fluid, whereby the most precise structure of the peptide, aptamer and complex can be obtained under physiological conditions. The NMR data can provide insights to the dynamic properties of the individual molecular structures and identify nucleotides involved in Watson-

crick base pairing. The studies can also provide structural, thermodynamic and kinetic aspects of the interactions present between the aptamer and the peptide. Furthermore, structural changes under the influence of potassium and sodium ion under varied concentrations can be explored [244,245]. Alternatively, X-ray crystallography studies can be an addition in gaining a better understanding of the chemical bonds, along with identifying the types of interaction that are present within the aptamer folding, peptide and the peptide-aptamer complex. X-ray studies aid in deducing the structure of the molecular complex by assessing the position of the atoms [246]. The structural folding and binding data would allow further optimisation of the aptamer sequence by removing bases which are not necessarily involved in the binding to the peptides

Aside from aptamer modification for improved affinity, the aptamer could be further functionalised with a different form of PEG, such as a four arm peg. If four aptamer molecules could be attached to one single polymer of PEG, then this would have the potential to significantly improve the efficacy of the aptamer for *in vivo* investigations.

Finally, as the MUC1/Y aptamers have shown binding specificity for the MUC1/Y protein, other than being employed as a direct therapeutic agent, they could be adapted for use as drug delivery agents. This typically involves activating a liposome that has encapsulated a therapeutic drug, by attaching a ligand, such as an aptamer, on its surface, which is capable of binding to a specific epitope present on the target site. The aptamer would bring an added advantage by increasing the therapeutic response due to delivering the drug directly to the target site, preventing the drug from being released in any healthy organs [247].

REFERENCE

1. Parkin, D. M., Bray, F., Ferlay, J. & Pisani, P. Global cancer statistics, 2002. *Ca-A Cancer Journal for Clinicians* **55**, 74-108 (2005).
2. www.cancer.org/.../Global_Cancer_Facts_and_Figures_2007_rev.pdf. Global Cancer Facts and Figures 2007-2008. *American Cancer Society* (2007).
3. Cerchia, L., Hamm, J., Libri, D., Tavitian, B. & De Franciscis, V. Nucleic acid aptamers in cancer medicine. *FEBS Letters* **528**, 12-16 (2002).
4. Devine, P. L. & McKenzie, I. F. C. Mucins: Structure, function, and associations with malignancy. *BioEssays* **14**, 619-625 (1992).
5. Bansil, R., Stanley, E. & LaMont, J. T. Mucin biophysics. *Annual Review of Physiology* **57**, 635-657 (1995).
6. Gendler, S. J. & Spicer, A. P. Epithelial mucin genes. *Annual Review of Physiology* **57**, 607-634 (1995).
7. Ferez-Vilar, J. & Hill, R. L. The structure and assembly of secreted mucins. *Journal of Biological Chemistry* **274**, 31751-31754 (1999).
8. Duraisamy, S., Ramasamy, S., Kharbanda, S. & Kufe, D. Distinct evolution of the human carcinoma-associated transmembrane mucins, MUC1, MUC4 AND MUC16. *Gene* **373**, 28-34 (2006).
9. Gendler, S. J. MUC1, the renaissance molecule. *Journal of mammary gland biology and neoplasia* **6**, 339-53 (2001).
10. Singh, P. K. & Hollingsworth, M. A. Cell surface-associated mucins in signal transduction. *Trends in Cell Biology* **16**, 467-476 (2006).
11. Brayman, M., Thathiah, A. & Carson, D. D. MUC1: A multifunctional cell surface component of reproductive tissue epithelia. *Reproductive Biology and Endocrinology* **2** (2004).
12. Patton, S., Gendler, S. J. & Spicer, A. P. The epithelial mucin, MUC1, of milk, mammary gland and other tissues. *Biochimica et Biophysica Acta - Reviews on Biomembranes* **1241**, 407-423 (1995).
13. Gendler, S. J. et al. Molecular cloning and expression of human tumor-associated polymorphic epithelial mucin. *Journal of Biological Chemistry* **265**, 15286-15293 (1990).
14. Hey, N. et al. Transmembrane and truncated (SEC) isoforms of MUC1 in the human endometrium and Fallopian tube. *Reproductive Biology and Endocrinology* **1** (2003).
15. Zrihan-Licht, S. et al. Characterization and molecular cloning of a novel MUC1 protein, devoid of tandem repeats, expressed in human breast cancer tissue. *European Journal of Biochemistry* **224**, 787-795 (1994).
16. Hilkens, J. et al. Is episialin/MUC1 involved in breast cancer progression? *Cancer Letters* **90**, 27-33 (1995).
17. Zrihan-Licht, S., Keydar, I., Elroy-Stein, O., Baruch, A. & Wreschner, D. H. Tyrosine phosphorylation of the MUC1 breast cancer membrane proteins Cytokine receptor-like molecules. *FEBS Letters* **356**, 130-136 (1994).
18. Parry, G., Beck, J. C., Moss, L., Bartley, J. & Ojakian, G. K. Determination of apical membrane polarity in mammary epithelial cell cultures: The role of cell-cell, cell-substratum, and membrane-cytoskeleton interactions. *Experimental Cell Research* **188**, 302-311 (1990).
19. Ligtenberg, M. J. L. et al. Cell-associated episialin is a complex containing two proteins derived from a common precursor. *Journal of Biological Chemistry* **267**, 6171-6177 (1992).
20. Julian, J. & Carson, D. D. Formation of MUC1 metabolic complex is conserved in tumor-derived and normal epithelial cells. *Biochemical and Biophysical Research Communications* **293**, 1183-1190 (2002).
21. Levitin, F. et al. The MUC1 SEA module is a self-cleaving domain. *Journal of Biological Chemistry* **280**, 33374-33386 (2005).

22. Litvinov, S. V. & Hilkens, J. The epithelial sialomucin, episialin, is sialylated during recycling. *Journal of Biological Chemistry* **268**, 21364-21371 (1993).
23. Parry, S. et al. Identification of MUC1 proteolytic cleavage sites in vivo. *Biochemical and Biophysical Research Communications* **283**, 715-720 (2001).
24. Burdick, M. D., Harris, A., Reid, C. J., Iwamura, T. & Hollingsworth, M. A. Oligosaccharides expressed on MUC1 produced by pancreatic and colon tumor cell lines. *Journal of Biological Chemistry* **272**, 24198-24202 (1997).
25. Ho, J. J. L., Cheng, S. & Kim, Y. S. Access to peptide regions of a surface mucin (MUC1) is reduced by sialic acids. *Biochemical and Biophysical Research Communications* **210**, 866-873 (1995).
26. Munster-Kuhnel, A. K. et al. Structure and function of vertebrate CMP-sialic acid synthetases. *Glycobiology* **14**, 43R-51R (2004).
27. Ugorski, M. & Laskowska, A. Sialyl Lewisa: A tumor-associated carbohydrate antigen involved in adhesion and metastatic potential of cancer cells. *Acta Biochimica Polonica* **49**, 303-311 (2002).
28. Dent, G. A., Civalier, C. J., Brecher, M. E. & Bentley, S. A. MUC1 expression in hematopoietic tissues. *American Journal of Clinical Pathology* **111**, 741-747 (1999).
29. Agrawal, B., Gendler, S. J. & Longenecker, B. M. The biological role of mucins in cellular interactions and immune regulation: Prospects for cancer immunotherapy. *Molecular Medicine Today* **4**, 397-403 (1998).
30. Agrawal, B., Krantz, M. J., Parker, J. & Longenecker, B. M. Expression of MUC1 mucin on activated human T cells: Implications for a role of MUC1 in normal immune regulation. *Cancer Research* **58**, 4079-4081 (1998).
31. Agrawal, B., Reddish, M. A. & Longenecker, B. M. In Vitro Induction of MUC-1 Peptide-Specific Type 1 T Lymphocyte and Cytotoxic T Lymphocyte Responses from Healthy Multiparous Donors. *Journal of Immunology* **157**, 2089-2095 (1996).
32. Pandey, P., Kharbanda, S. & Kufe, D. Association of the DF3/MUC1 breast cancer antigen with Grb2 and the Sos/Ras exchange protein. *Cancer Research* **55**, 4000-4003 (1995).
33. Schroeder, J. A., Thompson, M. C., Gardner, M. M. & Gendler, S. J. Transgenic MUC1 Interacts with Epidermal Growth Factor Receptor and Correlates with Mitogen-activated Protein Kinase Activation in the Mouse Mammary Gland. *Journal of Biological Chemistry* **276**, 13057-13064 (2001).
34. Yamamoto, M., Bharti, A., Li, Y. & Kufe, D. Interaction of the DF3/MUC1 breast carcinoma-associated antigen and β-catenin in cell adhesion. *Journal of Biological Chemistry* **272**, 12492-12494 (1997).
35. Li, Y., Kuwahara, H., Ren, J., Wen, G. & Kufe, D. The c-Src Tyrosine Kinase Regulates Signaling of the Human DF3/MUC1 Carcinoma-associated Antigen with GSK3β and β-Catenin. *Journal of Biological Chemistry* **276**, 6061-6064 (2001).
36. Wang, H., Lillehoj, E. P. & Kim, K. C. Identification of four sites of stimulated tyrosine phosphorylation in the MUC1 cytoplasmic tail. *Biochemical and Biophysical Research Communications* **310**, 341-346 (2003).
37. Baruch, A. et al. Preferential expression of novel MUC1 tumor antigen isoforms in human epithelial tumors and their tumor-potentiating function. *International Journal of Cancer* **71**, 741-749 (1997).
38. Bafna, S., Kaur, S. & Batra, S. K. Membrane-bound mucins: The mechanistic basis for alterations in the growth and survival of cancer cells. *Oncogene* **29**, 2893-2904 (2010).
39. Merlin, J. et al. Galectin-3 regulates MUC1 and EGFR cellular distribution and EGFR downstream pathways in pancreatic cancer cells. *Oncogene* **30**, 2514-2525 (2011).
40. Senapati, S., Das, S. & Batra, S. K. Mucin-interacting proteins: From function to therapeutics. *Trends in Biochemical Sciences* **35**, 236-245 (2010).

41. Bitter, B. G. et al. Intracellular MUC1 peptides inhibit cancer progression. *Clinical Cancer Research* **15**, 100-109 (2009).
42. Kohlgraf, K. G. et al. Contribution of the MUC1 tandem repeat and cytoplasmic tail to invasive and metastatic properties of a pancreatic cancer cell line. *Cancer Research* **63**, 5011-5020 (2003).
43. Engelmann, K. et al. Transmembrane and secreted MUC1 probes show trafficking-dependent changes in O-glycan core profiles. *Glycobiology* **15**, 1111-1124 (2005).
44. Hanisch, F.-G. & Muller, S. MUC1: The polymorphic appearance of a human mucin. *Glycobiology* **10**, 439-449 (2000).
45. Pudelko, M. et al. Formation of lactones from sialylated MUC1 glycopeptides. *Organic and Biomolecular Chemistry* **4**, 713-720 (2006).
46. Burchell, J. M., Mungul, A. & Taylor-Papadimitriou, J. O-linked glycosylation in the mammary gland: Changes that occur during malignancy. *Journal of Mammary Gland Biology and Neoplasia* **6**, 355-364 (2001).
47. Parry, S. et al. N-Glycosylation of the MUC1 mucin in epithelial cells and secretions. *Glycobiology* **16**, 623-634 (2006).
48. Dekker, J., Rossen, J. W. A., Buller, H. A. & Einerhand, A. W. C. The MUC family: An obituary. *Trends in Biochemical Sciences* **27**, 126-131 (2002).
49. Hilkens, J. & Buijs, F. Biosynthesis of MAM-6, an epithelial sialomucin. Evidence for involvement of a rare proteolytic cleavage step in the endoplasmic reticulum. *Journal of Biological Chemistry* **263**, 4215-4222 (1988).
50. Obermair, A. et al. Expression of MUC1 splice variants in benign and malignant ovarian tumours. *International Journal of Cancer* **100**, 166-171 (2002).
51. Baruch, A. et al. The breast cancer-associated MUC1 gene generates both a receptor and its cognate binding protein. *Cancer Research* **59**, 1552-1561 (1999).
52. Obermair, A. et al. Novel MUC1 Splice Variants Are Expressed in Cervical Carcinoma. *Gynecologic Oncology* **83**, 343-347 (2001).
53. Schut, I. C. et al. MUC1 expression, splice variant and short form transcription (MUC1/Z, MUC1/Y) in prostate cell lines and tissue. *BJU International* **91**, 278-283 (2003).
54. Passebosc-Faure, K. et al. Evaluation of a panel of molecular markers for the diagnosis of malignant serous effusions. *Clinical Cancer Research* **11**, 6862-6867 (2005).
55. Gilbert, J. C. et al. First-in-human evaluation of anti-von Willebrand factor therapeutic aptamer ARC1779 in healthy volunteers. *Circulation* **116**, 2678-2686 (2007).
56. Arzamendi, D. et al. An anti-von Willebrand factor aptamer reduces platelet adhesion among patients receiving aspirin and clopidogrel in an ex vivo shear-induced arterial thrombosis. *Clinical and Applied Thrombosis/Hemostasis* **17**, E70-E78 (2011).
57. Esposito, C. L., Catuogno, S., de Franciscis, V. & Cerchia, L. New insight into clinical development of nucleic acid aptamers. *Discovery medicine* **11**, 487-496 (2011).
58. Becker, R. C., Povsic, T., Cohen, M. G., Rusconi, C. P. & Sullenger, B. Nucleic acid aptamers as antithrombotic agents: Opportunities in extracellular therapeutics. *Thrombosis and Haemostasis* **103**, 586-595 (2010).
59. Wang, P. et al. Aptamers as therapeutics in cardiovascular diseases. *Current Medicinal Chemistry* **18**, 4169-4174 (2011).
60. Tuerk, C., MacDougal, S. & Gold, L. RNA pseudoknots that inhibit human immunodeficiency virus type 1 reverse transcriptase. *Proceedings of the National Academy of Sciences of the United States of America* **89**, 6988-6992 (1992).
61. Sung, H. J., Kayhan, B., Ben-Yedidia, T. & Arnon, R. A DNA aptamer prevents influenza infection by blocking the receptor binding region of the viral hemagglutinin. *Journal of Biological Chemistry* **279**, 48410-48419 (2004).

62. Charlton, J., Kirschenheuter, G. P. & Smith, D. Highly potent irreversible inhibitors of neutrophil elastase generated by selection from a randomized DNA-valine phosphonate library. *Biochemistry* **36**, 3018-3026 (1997).
63. Bless, N. M. et al. Protective effects of an aptamer inhibitor of neutrophil elastase in lung inflammatory injury. *Current Biology* **7**, 877-880 (1997).
64. Hicke, B. J. et al. DNA aptamers block L-selectin function in vivo. Inhibition of human lymphocyte trafficking in SCID mice. *Journal of Clinical Investigation* **98**, 2688-2692 (1996).
65. Watson, S. R. et al. Anti-L-selectin aptamers: Binding characteristics, pharmacokinetic parameters, and activity against an intravascular target in vivo. *Antisense and Nucleic Acid Drug Development* **10**, 63-75 (2000).
66. Ruckman, J. et al. 2 '-fluoropyrimidine RNA-based aptamers to the 165-amino acid form of vascular endothelial growth factor (VEGF(165)) - Inhibition of receptor binding and VEGF-induced vascular permeability through interactions requiring the exon 7-encoded domain. *Journal of Biological Chemistry* **273**, 20556-20567 (1998).
67. Guyer, D. R. et al. Anti-vascular endothelial growth factor therapy for subfoveal choroidal neovascularization secondary to age-related macular degeneration: Phase II study results. *Ophthalmology* **110**, 979-986 (2003).
68. Martin, D. F. et al. Preclinical and phase 1A clinical evaluation of an anti-VEGF pegylated aptamer (EYE001) for the treatment of exudative age-related macular degeneration. *Retina* **22**, 143-152 (2002).
69. Nieuwlandt, D., Wecker, M. & Gold, L. In vitro selection of RNA ligands to substance P. *Biochemistry* **34**, 5651-5659 (1995).
70. Binkley, J. et al. RNA ligands to human nerve growth factor. *Nucleic Acids Research* **23**, 3198-3205 (1995).
71. Soundararajan, S., Chen, W., Spicer, E. K., Courtenay-Luck, N. & Fernandes, D. J. The nucleolin targeting aptamer AS1411 destabilizes Bcl-2 messenger RNA in human breast cancer cells. *Cancer Research* **68**, 2358-2365 (2008).
72. Ireson, C. R. & Kelland, L. R. Discovery and development of anticancer aptamers. *Molecular Cancer Therapeutics* **5**, 2957-2962 (2006).
73. Daniels, D. A., Chen, H., Hicke, B. J., Swiderek, K. M. & Gold, L. A tenascin-C aptamer identified by tumor cell SELEX: Systematic evolution of ligands by exponential enrichment. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 15416-15421 (2003).
74. Hicke, B. J. et al. Tenascin-C Aptamers Are Generated Using Tumor Cells and Purified Protein. *Journal of Biological Chemistry* **276**, 48644-48654 (2001).
75. Lupold, S. E., Hicke, B. J., Lin, Y. & Coffey, D. S. Identification and characterization of nuclease-stabilized RNA molecules that bind human prostate cancer cells via the prostate-specific membrane antigen. *Cancer Research* **62**, 4029-4033 (2002).
76. Belsey, M. & Pavlou, A. K. The DNA/RNA market to 2010. *Journal of Commercial Biotechnology* **11**, 275 (2005).
77. Belikova, A. M., Zarytova, V. F. & Grineva, N. I. Synthesis of ribonucleosides and diribonucleoside phosphates containing 2-chloro-ethylamine and nitrogen mustard residues. *Tetrahedron Letters* **8**, 3557-3562 (1967).
78. Zamecnik, P. C. & Stephenson, M. L. Inhibition of Rous sarcoma virus replication and cell transformation by a specific oligodeoxynucleotide. *Proceedings of the National Academy of Sciences of the United States of America* **75**, 280-284 (1978).
79. Da Ros, T. et al. Oligonucleotides and oligonucleotide conjugates: A new approach for cancer treatment. *Current Medicinal Chemistry* **12**, 71-88 (2005).
80. Gallo, M., Montserrat, J. M. & Iribarren, A. M. Design and applications of modified oligonucleotides. *Brazilian Journal of Medical and Biological Research* **36**, 143-151 (2003).

81. Dias, N. & Stein, C. A. Antisense oligonucleotides: basic concepts and mechanisms. *Molecular cancer therapeutics* **1**, 347-355 (2002).
82. Chan, J. H. P., Lim, S. & Wong, W. S. F. Antisense oligonucleotides: From design to therapeutic application. *Clinical and Experimental Pharmacology and Physiology* **33**, 533-540 (2006).
83. Crooke, S. T. Vitravene® - Another piece in the mosaic. *Antisense and Nucleic Acid Drug Development* **8**, VII-VIII (1998).
84. Faria, M. & Ulrich, H. The use of synthetic oligonucleotides as protein inhibitors and anticode drugs in cancer therapy: Accomplishments and limitations. *Current Cancer Drug Targets* **2**, 355-368 (2002).
85. Letai, A. G., Palladino, M. A., Fromm, E., Rizzo, V. & Fresco, J. R. Specificity in formation of triple-stranded nucleic acid helical complexes: Studies with agarose-linked polyribonucleotide affinity columns. *Biochemistry* **27**, 9108-9112 (1988).
86. Jarald, E., Edwin, S., Dubey, P., Ajay Tiwari, A. & Thakre, V. Nucleic acid drugs: A novel approach. *African Journal of Biotechnology* **3**, 662-666 (2004).
87. Stull, R. A. & Szoka Jr., F. C. Antigene, ribozyme and aptamer nucleic acid drugs: Progress and prospects. *Pharmaceutical Research* **12**, 465-483 (1995).
88. Jimenez, R. M., Delwart, E. & LuptÃ¡k, A. Structure-based search reveals hammerhead ribozymes in the human microbiome. *Journal of Biological Chemistry* **286**, 7737-7743 (2011).
89. Forster, A. C. & Symons, R. H. Self-cleavage of virusoid RNA is performed by the proposed 55-nucleotide active site. *Cell* **50**, 9-16 (1987).
90. Hampel, A. & Tritz, R. RNA catalytic properties of the minimum (-)sTRSV sequence. *Biochemistry* **28**, 4929-4933 (1989).
91. Wu, H.-N. et al. Human hepatitis delta virus RNA subfragments contain an autocleavage activity. *Proceedings of the National Academy of Sciences of the United States of America* **86**, 1831-1835 (1989).
92. Sharmeen, L., Kuo, M. Y. P., Dinter-Gottlieb, G. & Taylor, J. Antigenomic RNA of human hepatitis delta virus can undergo self-cleavage. *Journal of Virology* **62**, 2674-2679 (1988).
93. Saville, B. J. & Collins, R. A. A site-specific self-cleavage reaction performed by a novel RNA in Neurospora mitochondria. *Cell* **61**, 685-696 (1990).
94. Winkler, W. C., Nahvi, A., Roth, A., Collins, J. A. & Breaker, R. R. Control of gene expression by a natural metabolite-responsive ribozyme. *Nature* **428**, 281-286 (2004).
95. Nielsen, H., Westhof, E. & Johansen, S. Molecular biology: An mRNA is capped by a 2',5' lariat catalyzed by a group I-like ribozyme. *Science* **309**, 1584-1587 (2005).
96. Nielsen, H. & Johansen, S. D. Group I introns: Moving in new directions. *RNA Biology* **6** (2009).
97. Prody, G. A., Bakos, J. T., Buzayan, J. M., Schneider, I. R. & Bruening, G. Autolytic processing of dimeric plant virus satellite RNA. *Science* **231**, 1577-1580 (1986).
98. Epstein, L. M. & Gall, J. G. Transcripts of newt satellite DNA self-cleave in vitro. *Cold Spring Harbor Symposia on Quantitative Biology* **52**, 261-265 (1987).
99. Rojas, A. A. et al. Hammerhead-mediated processing of satellite pDo500 family transcripts from Dolichopoda cave crickets. *Nucleic Acids Research* **28**, 4037-4043 (2000).
100. Ferbeyre, G., Smith, J. M. & Cedergren, R. Schistosome satellite DNA encodes active hammerhead ribozymes. *Molecular and Cellular Biology* **18**, 3880-3888 (1998).
101. Usman, N., Beigelman, L. & McSwiggen, J. A. Hammerhead ribozyme engineering. *Current Opinion in Structural Biology* **6**, 527-533 (1996).
102. Tedeschi, L., Lande, C., Cecchettini, A. & Citti, L. Hammerhead ribozymes in therapeutic target discovery and validation. *Drug Discovery Today* **14**, 776-783 (2009).

103. Buzayan, J. M., Gerlach, W. L. & Bruening, G. Non-enzymatic cleavage and ligation of RNAs complementary to a plant virus satellite RNA. *Nature* **323**, 349-353 (1986).
104. Buzayan, J. M., Hampel, A. & Bruening, G. Nucleotide sequence and newly formed phosphodiester bond of spontaneously ligated satellite tobacco ringspot virus RNA. *Nucleic Acids Research* **14**, 9729-9743 (1986).
105. Ferre-D'Amare, A. R. & Rupert, P. B. The hairpin ribozyme: From crystal structure to function. *Biochemical Society Transactions* **30**, 1105-1109 (2002).
106. Cottrell, J. W., Kuzmin, Y. I. & Fedor, M. J. Functional analysis of hairpin ribozyme active site architecture. *Journal of Biological Chemistry* **282**, 13498-13507 (2007).
107. Rupert, P. B. & Ferre-D'Amare, A. R. Crystal structure of a hairpin ribozyme-inhibitor complex with implications for catalysis. *Nature* **410**, 780-786 (2001).
108. Breaker, R. R. & Joyce, G. F. A DNA enzyme that cleaves RNA. *Chemistry and Biology* **1**, 223-229 (1994).
109. Santoro, S. W. & Joyce, G. F. Mechanism and utility of an RNA-cleaving DNA enzyme. *Biochemistry* **37**, 13330-13342 (1998).
110. Abdelgany, A., Ealing, J., Wood, M. & Beeson, D. Selective DNAzyme-mediated cleavage of AChR mutant transcripts by targeting the mutation site or through mismatches in the binding arm. *J RNAi Gene Silencing* **1**, 32 - 37 (2005).
111. Achenbach, J. C., Chiuman, W., Cruz, R. P. G. & Li, Y. DNAzymes: From creation in vitro to application in vivo. *Current Pharmaceutical Biotechnology* **5**, 321-336 (2004).
112. Khachigian, L. M. DNAzymes: Cutting a path to a new class of therapeutics. *Current Opinion in Molecular Therapeutics* **4**, 119-121 (2002).
113. Lu, Z.-X. et al. Effect of EBV LMP1 targeted DNAzymes on cell proliferation and apoptosis. *Cancer Gene Therapy* **12**, 647-654 (2005).
114. Fire, A. et al. Potent and specific genetic interference by double-stranded RNA in caenorhabditis elegans. *Nature* **391**, 806-811 (1998).
115. Napoli, C., Lemieux, C. & Jorgensen, R. Introduction of a chimeric chalcone synthase gene into petunia results in reversible co-suppression of homologous genes in trans. *Plant Cell* **2**, 279-289 (1990).
116. Cogoni, C. & Macino, G. Posttranscriptional gene silencing in Neurospora by a RecQ DNA helicase. *Science* **286**, 2342-2344 (1999).
117. Elbashir, S. M. et al. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* **411**, 494-498 (2001).
118. Seiffert, S. et al. Characterization of side reactions during the annealing of small interfering RNAs. *Analytical Biochemistry* (2011).
119. Tuschl, T. RNA interference and small interfering RNAs. *ChemBioChem* **2**, 239-245 (2001).
120. Fichou, Y. & Ferec, C. The potential of oligonucleotides for therapeutic applications. *Trends in Biotechnology* **24**, 563-570 (2006).
121. Micura, R. Small interfering RNAs and their chemical synthesis. *Angewandte Chemie - International Edition* **41**, 2265-2269 (2002).
122. Meister, G. & Tuschl, T. Mechanisms of gene silencing by double-stranded RNA. *Nature* **431**, 343-349 (2004).
123. Jayasena, S. D. Aptamers: An emerging class of molecules that rival antibodies in diagnostics. *Clinical Chemistry* **45**, 1628-1650 (1999).
124. Nimjee, S. M., Rusconi, C. P. & Sullenger, B. A. *Aptamers: An emerging class of therapeutics* (2005).
125. Ellington, A. D. & Szostak, J. W. In vitro selection of RNA molecules that bind specific ligands. *Nature* **346**, 818-822 (1990).
126. Rimmele, M. Nucleic acid aptamers as tools and drugs: Recent developments. *ChemBioChem* **4**, 963-971 (2003).

127. Stoltenburg, R., Reinemann, C. & Strehlitz, B. SELEX--a (r)evolutionary method to generate high-affinity nucleic acid ligands. *Biomol Eng* **24**, 381-403 (2007).
128. Boiziau, C., Dausse, E., Yurchenko, L. & Toulme, J.-J. DNA aptamers selected against the HIV-1 trans-activation-responsive RNA element form RNA-DNA kissing complexes. *Journal of Biological Chemistry* **274**, 12730-12737 (1999).
129. Huizenga, D. E. & Szostak, J. W. A DNA aptamer that binds adenosine and ATP. *Biochemistry* **34**, 656-665 (1995).
130. Sassanfar, M. & Szostak, J. W. An RNA motif that binds ATP. *Nature* **364**, 550-553 (1993).
131. Haller, A. A. & Sarnow, P. In vitro selection of a 7-methyl-guanosine binding RNA that inhibits translation of capped mRNA molecules. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 8521-8526 (1997).
132. Green, L. S. et al. Inhibitory DNA ligands to platelet-derived growth factor B-chain. *Biochemistry* **35**, 14413-14424 (1996).
133. Jellinek, D., Green, L.S., Bell, C., Janjic, N. Inhibition of receptor binding by high-affinity RNA ligands to vascular endothelial growth factor. *Biochemistry* **33**, 10450-10456 (1994).
134. Giver, L., Bartel, D. P., Zapp, M. L., Green, M. R. & Ellington, A. D. Selection and design of high-affinity RNA ligands for HIV-1 Rev. *Gene* **137**, 19-24 (1993).
135. Giver, L. et al. Selective optimization of the Rev-binding element of HIV-1. *Nucleic Acids Research* **21**, 5509-5516 (1993).
136. Harada, K. & Frankel, A. D. Identification of two novel arginine binding DNAs. *EMBO Journal* **14**, 5798-5811 (1995).
137. Connell, G. J., Illangesekare, M. & Yarus, M. Three small ribooligonucleotides with specific arginine sites. *Biochemistry* **32**, 5497-5502 (1993).
138. Wilson, C. & Szostak, J. W. Isolation of a fluorophore-specific DNA aptamer with weak redox activity. *Chemistry and Biology* **5**, 609-617 (1998).
139. Ellington, A. D. & Szostak, J. W. Selection in vitro of single-stranded DNA molecules that fold into specific ligand-binding structures. *Nature* **355**, 850-852 (1992).
140. Hofmann, H.-P., Limmer, S., Hornung, V. & Sprinzl, M. Ni^{2+} binding RNA motifs with an asymmetric purine-rich internal loop and a G-A base pair. *RNA* **3**, 1289-1300 (1997).
141. Ciesiolka, J., Gorski, J. & Yarus, M. Selection of an RNA domain that binds Zn^{2+}. *RNA* **1**, 538-550 (1995).
142. Shangguan, D. et al. Aptamers evolved from live cells as effective molecular probes for cancer study. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 11838-11843 (2006).
143. Blank, M., Weinschenk, T., Priemer, M. & Schluesener, H. Systematic evolution of a DNA aptamer binding to rat brain tumor microvessels: Selective targeting of endothelial regulatory protein pigpen. *Journal of Biological Chemistry* **276**, 16464-16468 (2001).
144. Lorger, M., Engstler, M., Homann, M. & Goringer, H. U. Targeting the variable surface of African trypanosomes with variant surface glycoprotein-specific, serum-stable RNA aptamers. *Eukaryotic Cell* **2**, 84-94 (2003).
145. Feigon, J., Dieckmann, T. & Smith, F. W. Aptamer structures from A to zeta. *Chemistry & Biology* **3**, 611-617 (1996).
146. Gold, L., Polisky, B., Uhlenbeck, O. & Yarus, M. Diversity of Oligonucleotide Functions. *Annual Review of Biochemistry* **64**, 763-797 (1995).
147. Mok, W. & Li, Y. Recent progress in nucleic acid aptamer-based biosensors and bioassays. *Sensors* **8**, 7050-7084 (2008).
148. Nolte, A., Klussmann, S., Bald, R., Erdmann, V. A. & Furste, J. P. Mirror-design of L-oligonucleotide ligands binding to L-arginine. *Nature Biotechnology* **14**, 1116-1119 (1996).

149. Klussmann, S., Nolte, A., Bald, R., Erdmann, V. A. & Furste, J. P. Mirror-image RNA that binds D-adenosine. *Nature Biotechnology* **14**, 1112-1115 (1996).
150. Purschke, W. G., Radtke, F., Kleinjung, F. & Klussmann, S. A DNA Spiegelmer to staphylococcal enterotoxin B. *Nucleic Acids Research* **31**, 3027-3032 (2003).
151. Eulberg, D. & Klussmann, S. Spiegelmers: Biostable aptamers. *ChemBioChem* **4**, 979-983 (2003).
152. Tuerk, C. & Gold, L. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* **249**, 505-510 (1990).
153. Luzi, E., Minunni, M., Tombelli, S. & Mascini, M. New trends in affinity sensing: Aptamers for ligand binding. *TrAC - Trends in Analytical Chemistry* **22**, 810-818 (2003).
154. White, R. R., Sullenger, B. A. & Rusconi, C. P. Developing aptamers into therapeutics. *Journal of Clinical Investigation* **106**, 929-934 (2000).
155. Hicke, B. J. & Stephens, A. W. Escort aptamers: a delivery service for diagnosis and therapy. *Journal of Clinical Investigation* **106**, 923-928 (2000).
156. Zhou, B. & Wang, B. Pegaptanib for the treatment of age-related macular degeneration. *Experimental Eye Research* **83**, 615-619 (2006).
157. Ng, E. W. M. et al. Pegaptanib, a targeted anti-VEGF aptamer for ocular vascular disease. *Nature Reviews Drug Discovery* **5**, 123-132 (2006).
158. Lee, J.-H. et al. A therapeutic aptamer inhibits angiogenesis by specifically targeting the heparin binding domain of VEGF165. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 18902-18907 (2005).
159. Gatto, B. & Cavalli, M. From proteins to nucleic acid-based drugs: The role of biotech in anti-VEGF therapy. *Anti-Cancer Agents in Medicinal Chemistry* **6**, 287-301 (2006).
160. Huang, J. et al. Highly specific antiangiogenic therapy is effective in suppressing growth of experimental Wilms tumors. *Journal of Pediatric Surgery* **36**, 357-361 (2001).
161. Xu, X. et al. Inhibition of DNA Replication and Induction of S Phase Cell Cycle Arrest by G-rich Oligonucleotides. *Journal of Biological Chemistry* **276**, 43221-43230 (2001).
162. Bates, P. J., Kahlon, J. B., Thomas, S. D., Trent, J. O. & Miller, D. M. Antiproliferative activity of G-rich oligonucleotides correlates with protein binding. *Journal of Biological Chemistry* **274**, 26369-26377 (1999).
163. Dapic, V. et al. Antiproliferative activity of G-quartet-forming oligonucleotides with backbone and sugar modifications. *Biochemistry* **41**, 3676-3685 (2002).
164. http://www.clinicaltrials.gov/.
165. Silver, D. A., Pellicer, I., Fair, W. R., Heston, W. D. W. & Cordon-Cardo, C. Prostate-specific membrane antigen expression in normal and malignant human tissues. *Clinical Cancer Research* **3**, 81-85 (1997).
166. Liu, C. et al. Prostate-specific membrane antigen directed selective thrombotic infarction of tumors. *Cancer Research* **62**, 5470-5475 (2002).
167. Rajasekaran, A. K., Anilkumar, G. & Christiansen, J. J. Is prostate-specific membrane antigen a multifunctional protein? *American Journal of Physiology - Cell Physiology* **288**, C975-C981 (2005).
168. Farokhzad, O. C. et al. Targeted nanoparticle-aptamer bioconjugates for cancer chemotherapy in vivo. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 6315-6320 (2006).
169. Cheng, J. et al. Formulation of functionalized PLGA-PEG nanoparticles for in vivo targeted drug delivery. *Biomaterials* **28**, 869-876 (2007).
170. Pendergrast, P. S., Marsh, H. N., Grate, D., Healy, J. M. & Stanton, M. Nucleic acid aptamers for target validation and therapeutic applications. *Journal of biomolecular techniques : JBT.* **16**, 224-234 (2005).
171. Pietras, K. et al. Inhibition of platelet-derived growth factor receptors reduces interstitial hypertension and increases transcapillary transport in tumors. *Cancer Research* **61**, 2929-2934 (2001).

172. Pietras, K. et al. Inhibition of PDGF receptor signaling in tumor stroma enhances antitumor effect of chemotherapy. *Cancer Research* **62**, 5476-5484 (2002).
173. Ruegg, C. R., Chiquet-Ehrismann, R. & Alkan, S. S. Tenascin, an extracellular matrix protein, exerts immunomodulatory activities. *Proceedings of the National Academy of Sciences of the United States of America* **86**, 7437-7441 (1989).
174. Schmidt, K. S. et al. Application of locked nucleic acids to improve aptamer in vivo stability and targeting function. *Nucleic Acids Research* **32**, 5757-5765 (2004).
175. Hicke, B. J. et al. Tumor targeting by an aptamer. *Journal of Nuclear Medicine* **47**, 668-678 (2006).
176. Hesselberth, J., Robertson, M. P., Jhaveri, S. & Ellington, A. D. In vitro selection of nucleic acids for diagnostic applications. *Reviews in Molecular Biotechnology* **74**, 15-25 (2000).
177. Charlton, J., Sennello, J. & Smith, D. In vivo imaging of inflammation using an aptamer inhibitor of human neutrophil elastase. *Chemistry and Biology* **4**, 809-816 (1997).
178. Pieken, W. A., Olsen, D. B., Benseler, F., Aurup, H. & Eckstein, F. Kinetic Characterization of Ribonuclease-Resistant 2'-Modified Hammerhead Ribozymes. *Science* **253**, 314-317 (1991).
179. Blank, M. & Blind, M. Aptamers as tools for target validation. *Current Opinion in Chemical Biology* **9**, 336-342 (2005).
180. O'Sullivan, C. K. Aptasensors - The future of biosensing? *Fresenius' Journal of Analytical Chemistry* **372**, 44-48 (2002).
181. Wochner, A., Menger, M. & Rimmele, M. Characterisation of aptamers for therapeutic studies. *Expert Opinion on Drug Discovery* **2**, 1205-1224 (2007).
182. Brody, E. N. & Gold, L. Aptamers as therapeutic and diagnostic agents. *Reviews in Molecular Biotechnology* **74**, 5-13 (2000).
183. Ulrich, H. DNA and RNA aptamers as modulators of protein function. *Medicinal chemistry* **1**, 199-208 (2005).
184. Agrawal, S. Importance of nucleotide sequence and chemical modifications of antisense oligonucleotides. *Biochimica et Biophysica Acta - Gene Structure and Expression* **1489**, 53-67 (1999).
185. Younes, C. K., Boisgard, R. & Tavitian, B. Labelled oligonucleotides as radiopharmaceuticals: Pitfalls, problems and perspectives. *Current Pharmaceutical Design* **8**, 1451-1466 (2002).
186. Miller, P. S. Oligonucleoside methylphosphonates as antisense reagents. *Bio/Technology* **9**, 358-362 (1991).
187. Chem, T.-L., Miller, P. S., Ts'o, P. O. P. & Colvin, O. M. Disposition and metabolism of oligodeoxynucleoside methylphosphonate following a single iv injection in mice. *Drug Metabolism and Disposition* **18**, 815-818 (1990).
188. Eaton, B. E. The joys of in vitro selection: chemically dressing oligonucleotides to satiate protein targets. *Current Opinion in Chemical Biology* **1**, 10-16 (1997).
189. Booth, J. et al. Determining the origin of the stabilization of DNA by 5-aminopropynylation of pyrimidines. *Biochemistry* **44**, 4710-4719 (2005).
190. Abuchowski, A., McCoy, J. R. & Palczuk, N. C. Effect of covalent attachment of polyethylene glycol on immunogenicity and circulating life of bovine liver catalase. *Journal of Biological Chemistry* **252**, 3582-3586 (1977).
191. Zalipsky, S. Chemistry of polyethylene glycol conjugates with biologically active molecules. *Advanced Drug Delivery Reviews* **16**, 157-182 (1995).
192. Veronese, F. M. & Pasut, G. PEGylation, successful approach to drug delivery. *Drug Discovery Today* **10**, 1451-1458 (2005).
193. Roberts, M. J., Bentley, M. D. & Harris, J. M. Chemistry for peptide and protein PEGylation. *Advanced Drug Delivery Reviews* **54**, 459-476 (2002).

194. Bailon, P. & Berthold, W. Polyethylene glycol-conjugated pharmaceutical proteins. *Pharmaceutical Science & Technology Today* **1**, 352-356 (1998).
195. Harris, J. M. & Chess, R. B. Effect of pegylation on pharmaceuticals. *Nature Reviews Drug Discovery* **2**, 214-221 (2003).
196. Milton Harris, J., Martin, N. E. & Modi, M. Pegylation: A novel process for modifying pharmacokinetics. *Clinical Pharmacokinetics* **40**, 539-551 (2001).
197. Ferreira, C. S. M., Matthews, C. S. & Missailidis, S. DNA aptamers that bind to MUC1 tumour marker: Design and characterization of MUC1-binding single-stranded DNA aptamers. *Tumor Biology* **27**, 289-301 (2006).
198. Liss, M., Petersen, B., Wolf, H. & Prohaska, E. An aptamer-based quartz crystal protein biosensor. *Analytical Chemistry* **74**, 4488-4495 (2002).
199. Nitsche, A. et al. One-step selection of Vaccinia virus-binding DNA aptamers by MonoLEX. *BMC Biotechnol* **7**, 48-59 (2007).
200. Tang, Z. et al. Selection of aptamers for molecular recognition and characterization of cancer cells. *Analytical Chemistry* **79**, 4900-4907 (2007).
201. Pan, W., Xin, P. & Clawson, G. A. Minimal primer and primer-free SELEX protocols for selection of aptamers from random DNA libraries. *Biotechniques* **44**, 351-60 (2008).
202. Mendonsa, S. D. & Bowser, M. T. In Vitro Evolution of Functional DNA Using Capillary Electrophoresis. *Journal of the American Chemical Society* **126**, 20-21 (2004).
203. Mendonsa, S. D. & Bowser, M. T. In vitro selection of high-affinity DNA ligands for human IgE using capillary electrophoresis. *Analytical Chemistry* **76**, 5387-5392 (2004).
204. Berezovski, M. et al. Nonequilibrium capillary electrophoresis of equilibrium mixtures: A universal tool for development of aptamers. *Journal of the American Chemical Society* **127**, 3165-3171 (2005).
205. Berezovski, M., Musheev, M., Drabovich, A. & Krylov, S. N. Non-SELEX selection of aptamers. *Journal of the American Chemical Society* **128**, 1410-1411 (2006).
206. Fan, M. et al. Aptamer selection express: a novel method for rapid single-step selection and sensing of aptamers. *Journal of biomolecular techniques : JBT* **19**, 311-319 (2008).
207. Smith, D., Kirschenheuter, G. P., Charlton, J., Guidot, D. M. & Repine, J. E. In vitro selection of RNA-based irreversible inhibitors of human neutrophil elastase. *Chemistry and Biology* **2**, 741-750 (1995).
208. Yao, W., Adelman, K. & Bruenn, J. A. In vitro selection of packaging sites in a double-stranded RNA virus. *Journal of Virology* **71**, 2157-2162 (1997).
209. Tsai, R. Y. L. & Reed, R. R. Identification of DNA recognition sequences and protein interaction domains of the multiple-Zn-finger protein Roaz. *Molecular and Cellular Biology* **18**, 6447-6456 (1998).
210. Goodman, S. D., Velten, N. J., Gao, Q., Robinson, S. & Segall, A. M. In vitro selection of integration host factor binding sites. *Journal of Bacteriology* **181**, 3246-3255 (1999).
211. Park, S., Myszka, D. G., Yu, M., Littler, S. J. & Laird-Offringa, I. A. HuD RNA recognition motifs play distinct roles in the formation of a stable complex with AU-rich RNA. *Molecular and Cellular Biology* **20**, 4765-4772 (2000).
212. Katsamba, P. S., Myszka, D. G. & Laird-Offringa, I. A. Two Functionally Distinct Steps Mediate High Affinity Binding of U1A Protein to U1 Hairpin II RNA. *Journal of Biological Chemistry* **276**, 21476-21481 (2001).
213. Misono, T. S. & Kumar, P. K. R. Selection of RNA aptamers against human influenza virus hemagglutinin using surface plasmon resonance. *Analytical Biochemistry* **342**, 312-317 (2005).
214. Khati, M. et al. Neutralization of Infectivity of Diverse R5 Clinical Isolates of Human Immunodeficiency Virus Type 1 by gp120-Binding 2'F-RNA Aptamers. *Journal of Virology* **77**, 12692-12698 (2003).
215. Collett, J. R. et al. Functional RNA microarrays for high-throughput screening of antiprotein aptamers. *Analytical Biochemistry* **338**, 113-123 (2005).

216. Cho, E. J., Collett, J. R., Szafranska, A. E. & Ellington, A. D. Optimization of aptamer microarray technology for multiple protein targets. *Analytica Chimica Acta* **564**, 82-90 (2006).
217. Collett, J. R., Eun, J. C. & Ellington, A. D. Production and processing of aptamer microarrays. *Methods* **37**, 4-15 (2005).
218. Holmberg, A. et al. The biotin-streptavidin interaction can be reversibly broken using water at elevated temperatures. *Electrophoresis* **26**, 501-510 (2005).
219. Bayer, E. A., Ben-Hur, H., Gitlin, G. & Wilchek, M. An improved method for the single-step purification of streptavidin. *Journal of Biochemical and Biophysical Methods* **13**, 103-112 (1986).
220. Zuker, M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research* **31**, 3406-3415 (2003).
221. Healy, J. M. et al. Pharmacokinetics and biodistribution of novel aptamer compositions. *Pharmaceutical Research* **21**, 2234-2246 (2004).
222. Ruggiero, A. et al. Paradoxical glomerular filtration of carbon nanotubes. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 12369-12374 (2010).
223. Yamaoka, T., Tabata, Y. & Ikada, Y. Distribution and tissue uptake of poly(ethylene glycol) with different molecular weights after intravenous administration to mice. *Journal of Pharmaceutical Sciences* **83**, 601-606 (1994).
224. Veronese, F. M. Peptide and protein PEGylation: a review of problems and solutions. *Biomaterials* **22**, 405-417 (2001).
225. Floege, J. et al. Novel approach to specific growth factor inhibition in vivo: Antagonism of platelet-derived growth factor in glomerulonephritis by aptamers. *American Journal of Pathology* **154**, 169-179 (1999).
226. Kang, H., Park, J. K., Seu, Y.-B. & Hahn, S. K. *A novel branch-type PEGylation of aptamer therapeutics* (2007).
227. Boomer, R. M. et al. Conjugation to polyethylene glycol polymer promotes aptamer biodistribution to healthy and inflamed tissues. *Oligonucleotides* **15**, 183-195 (2005).
228. Rajender Reddy, K., Modi, M. W. & Pedder, S. Use of peginterferon alfa-2a (40 KD) (PegasysÂ®) for the treatment of hepatitis C. *Advanced Drug Delivery Reviews* **54**, 571-586 (2002).
229. Bailon, P. et al. Rational design of a potent, long-lasting form of interferon: A 40 kDa branched polyethylene glycol-conjugated interferon alpha-2a for the treatment of hepatitis C. *Bioconjugate Chemistry* **12**, 195-202 (2001).
230. Skoog, B. Determination of polyethylene glycols 4000 and 6000 in plasma protein preparations. *Vox Sanguinis* **37**, 345-349 (1979).
231. Shaunak, S. et al. Site-specific PEGylation of native disulfide bonds in therapeutic proteins. *Nature Chemical Biology* **2**, 312-313 (2006).
232. Na, D. H. et al. Sodium dodecyl sulfate-capillary gel electrophoresis of polyethylene glycolylated interferon alpha. *Electrophoresis* **25**, 476-479 (2004).
233. Zhu, J. et al. Identification of ssDNA aptamers specific for anti-neuroexcitation peptide III and molecular modeling studies: Insights into structural interactions. *Archives of Pharmacal Research* **31**, 1120-1128 (2008).
234. Pileur, F. et al. Selective inhibitory DNA aptamers of the human RNase H1. *Nucleic Acids Research* **31**, 5776-5788 (2003).
235. Ferreira, C. S. M., Cheung, M. C., Missailidis, S., Bisland, S. & Gariepy, J. Phototoxic aptamers selectively enter and kill epithelial cancer cells. *Nucleic Acids Research* **37**, 866-876 (2009).
236. Shi, H., Fan, X., Sevilimedu, A. & Lis, J. T. RNA aptamers directed to discrete functional sites on a single protein structural domain. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 3742-3746 (2007).

237. Yang, C. et al. RNA aptamers targeting the cell death inhibitor CED-9 induce cell killing in Caenorhabditis elegans. *Journal of Biological Chemistry* **281**, 9137-9144 (2006).
238. Sainathan, S. K. et al. PEGylated murine Granulocyte-macrophage colony-stimulating factor: Production, purification, and characterization. *Protein Expression and Purification* **44**, 94-103 (2005).
239. Hwang, J. & Nishikawa, S. Novel approach to analyzing RNA aptamer-protein interactions: Toward further applications of aptamers. *Journal of Biomolecular Screening* **11**, 599-605 (2006).
240. Tang, Q., Su, X. & Loh, K. P. Surface plasmon resonance spectroscopy study of interfacial binding of thrombin to antithrombin DNA aptamers. *Journal of Colloid and Interface Science* **315**, 99-106 (2007).
241. Van Den Bossche, J. et al. Efficient receptor-independent intracellular translocation of aptamers mediated by conjugation to carbon nanotubes. *Chemical Communications* **46**, 7379-7381 (2010).
242. Cho, M. et al. Quantitative selection of DNA aptamers through microfluidic selection and high-throughput sequencing. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 15373-15378 (2010).
243. Nagarkatti, R. et al. Development of an aptamer-based concentration method for the detection of trypanosoma cruzi in blood. *PLoS ONE* **7** (2012).
244. Wuthrich, K. NMR studies of structure and function of biological macromolecules (Nobel Lecture). *Angewandte Chemie - International Edition* **42**, 3340-3363 (2003).
245. Robertson, S. A., Harada, K., Frankel, A. D. & Wemmer, D. E. Structure determination and binding kinetics of a DNA aptamer - Argininamide complex. *Biochemistry* **39**, 946-954 (2000).
246. Kay, L. E. NMR studies of protein structure and dynamics. *Journal of Magnetic Resonance* **173**, 193-207 (2005).
247. Forssen, E. & Willis, M. Ligand-targeted liposomes. *Advanced Drug Delivery Reviews* **29**, 249-271 (1998).

APPENDIX

CHAPTER 3.0

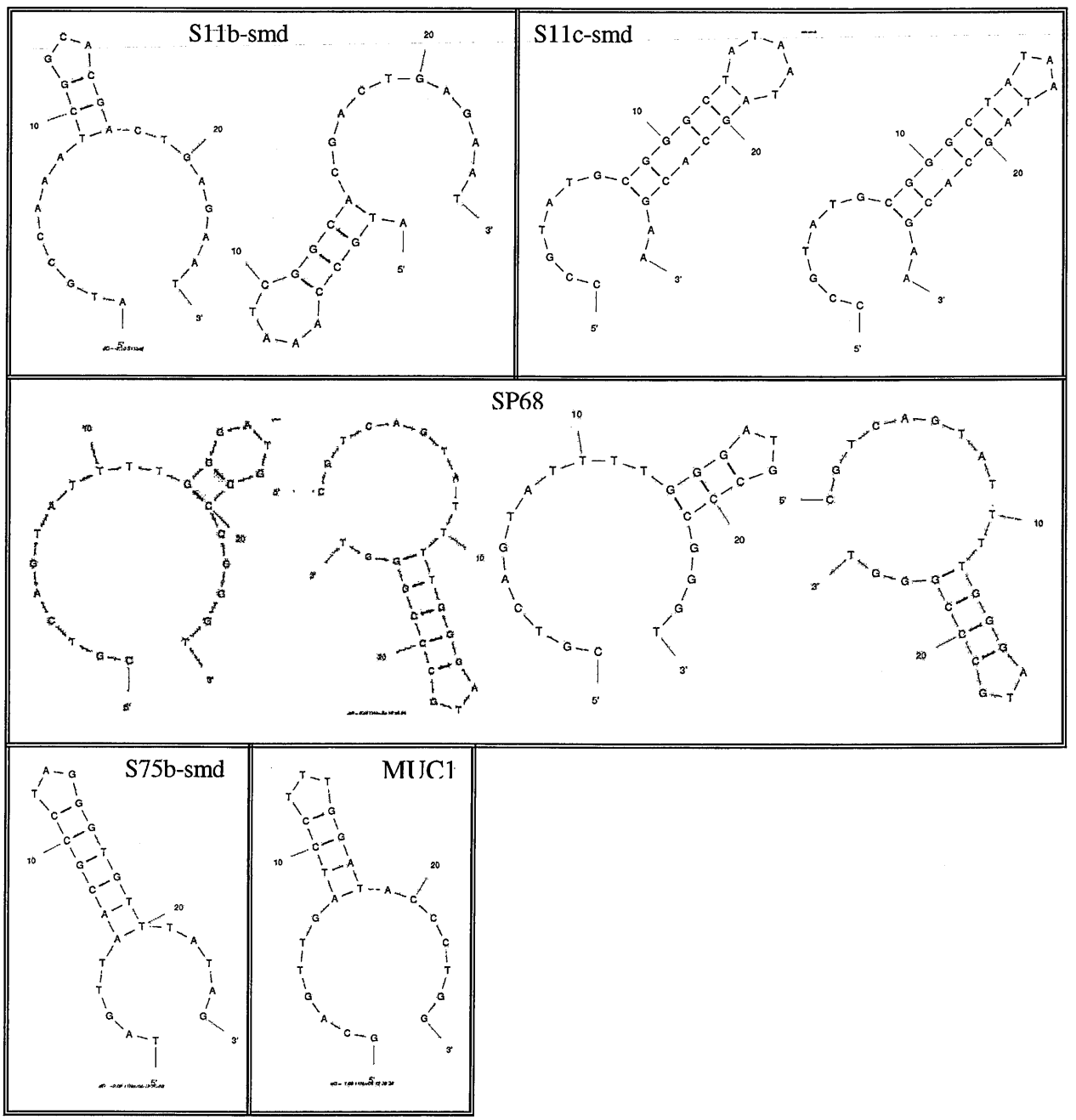
The forty aptamer sequences identifying the unknown variable region from the 1.5 M NaCl concentration and the 95°C samples for both MUC1/Y 10mer and 20mer peptides.

Aptamer sequences selected against the 10mer peptide at 1.5 M NaCl			
Name of Sequence	**Forward hybridising primer**	**Variable region**	**Reverse hybridising primer**
S11	GGGAGACAAGAATAAACGCTCAA	GGCAACATACTGTAAAGCTCAGGAC	TTCGACAGGAGGCTCACAACAGGC
S12	GGGAGACAAGAATAAACGCTCAA	CTAAAGTGCCTCACGCTGTTAACTC	TTCGACAGGAGGCTCACAACAGGC
S13	GGGAGACAAGAATAAACGCTCAA	TTATCACCTTAACTTCGTCGGACG	TTCGACAGGAGGCTCACAACAGGC
S15	GGGAGACAAGAATAAACGCTCAA	TTACTTGCTCAACACACGCCAGGAG	TTCGACAGGAGGCTCACAACAGGC
S18	GGGAGACAAGAATAAACGCTCAA	CGCCTAGTAAGATTTGACGGACCCC	TTCGACAGGAGGCTCACAACAGGC
S75	GGGAGACAAGAATAAACGCTCAA	CTGGTATCTATATGAAGGTTGTAGC	TTCGACAGGAGGCTCACAACAGGC
S76	GGGAGACAAGAATAAACGCTCAA	AAACCGAATAGAATCGCCTGACGGG	TTCGACAGGAGGCTCACAACAGGC
S85	GGGAGACAAGAATAAACGCTCAA	GATCTATCCTAACCCCGGTAGCGCC	TTCGACAGGAGGCTCACAACAGGC
S86	GGGAGACAAGAATAAACGCTCAA	CACTATGGCCACTCGACAGGAGGCT	TTCGACAGGAGGCTCACAACAGGC
S87	GGGAGACAAGAATAAACGCTCAA	GGCATTTTGTCCCTTGTTTGAACAA	TTCGACAGGAGGCTCACAACAGGC
Aptamer sequences selected against the 10mer peptide at 95°C			
Name of Sequence	**Forward hybridising primer**	**Variable region**	**Reverse hybridising primer**
S37	GGGAGACAAGAATAAACGCTCAA	TAACCCGACATCAACGCAATCTCAC	TTCGACAGGAGGCTCACAACAGGC
S38	GGGAGACAAGAATAAACGCTCAA	TCTCCCCGTGGGGTATGGGTGGAAA	TTCGACAGGAGGCTCACAACAGGC
S51	GGGAGACAAGAATAAACGCTCAA	AGATAAAAAGGCTGTCTGAAAATT	TTCGACAGGAGGCTCACAACAGGC
S52	GGGAGACAAGAATAAACGCTCAA	ATTAACTTCCAAGCTCTTGGACTAT	TTCGACAGGAGGCTCACAACAGGC
S56	GGGAGACAAGAATAAACGCTCAA	CCATCGGCGTCTAATACGTCA	TTCGACAGGAGGCTCACAACAGGC
S58	GGGAGACAAGAATAAACGCTCAA	GTAAGTTCTTAGTGCTGCCCTTTCA	TTCGACAGGAGGCTCACAACAGGC
S101	GGGAGACAAGAATAAACGCTCAA	TCACATAATAACCCGCAGCTAAAGT	TTCGACAGGAGGCTCACAACAGGC
S102	GGGAGACAAGAATAAACGCTCAA	AAATTTTATGGAGTTCTTTGGATAA	TTCGACAGGAGGCTCACAACAGGC
S114	GGGAGACAAGAATAAACGCTCAA	ACTTCACAAACGGTGCGGAATCCGA	TTCGACAGGAGGCTCACAACAGGC
S116	GGGAGACAAGAATAAACGCTCAA	ATCCTGTTATCTCCTTCTGCT	TTCGACAGGAGGCTCACAACAGGC

Aptamer sequences selected against the 20mer peptide at 1.5 M NaCl			
Name of Sequence	**Forward hybridising primer**	**Variable region**	**Reverse hybridising primer**
S6	GGGAGACAAGAATAAACGCTCAA	GCCAAGCATGCGTCACATGGCTGTG	TTCGACAGGAGGCTCACAACAGGC
S8	GGGAGACAAGAATAAACGCTCAA	TGAGTGGCACATGGCAGTACGTAGG	TTCGACAGGAGGCTCACAACAGGC
S9	GGGAGACAAGAATAAACGCTCAA	CGCCACCACCAGTCATAGCCCAGTA	TTCGACAGGAGGCTCACAACAGGC
S10	GGGAGACAAGAATAAACGCTCAA	TCGCTGACCATAGGTTCCTTAACTA	TTCGACAGGAGGCTCACAACAGGC
S31	GGGAGACAAGAATAAACGCTCAA	AATAACCTCAGCAGATCCAACGGGA	TTCGACAGGAGGCTCACAACAGGC
S32	GGGAGACAAGAATAAACGCTCAA	CGGCCCGTACGGCCA	TTCGACAGGAGGCTCACAACAGGC
S33	GGGAGACAAGAATAAACGCTCAA	TGGATAAACACCATTCACGGTAGGC	TTCGACAGGAGGCTCACAACAGGC
S35	GGGAGACAAGAATAAACGCTCAA	CTAATGCCCATATTCCAAGACACCG	TTCGACAGGAGGCTCACAACAGGC
S36	GGGAGACAAGAATAAACGCTCAA	TCAGCAATAGCCATCCATACTGGC	TTCGACAGGAGGCTCACAACAGGC
S78	GGGAGACAAGAATAAACGCTCAA	GCGCACGACTA	TTCGACAGGAGGCTCACAACAGGC
Aptamer sequences selected against the 20mer peptide at 95°C			
Name of Sequence	**Forward hybridising primer**	**Variable region**	**Reverse hybridising primer**
S41	GGGAGACAAGAATAAACGCTCAA	TCCATCTATAAATTCAAACAACATT	TTCGACAGGAGGCTCACAACAGGC
S44	GGGAGACAAGAATAAACGCTCAA	TGGCCGGGCGTTCCCCGTACAGAAC	TTCGACAGGAGGCTCACAACAGGC
S46	GGGAGACAAGAATAAACGCTCAA	ACAACACGCTAAAACGGACTGTTCT	TTCGACAGGAGGCTCACAACAGGC
S47	GGGAGACAAGAATAAACGCTCAA	TGAGGCTTCAAACTCTTATGTTCAG	TTCGACAGGAGGCTCACAACAGGC
S49	GGGAGACAAGAATAAACGCTCAA	GTAAAATTTAATTACATTTAATA	TTCGACAGGAGGCTCACAACAGGC
S88	GGGAGACAAGAATAAACGCTCAA	TCGGTAAAATACGACAGCCCGCAAA	TTCGACAGGAGGCTCACAACAGGC
S89	GGGAGACAAGAATAAACGCTCAA	TTGTCCCATACCGCATGCGTCTAGA	TTCGACAGGAGGCTCACAACAGGC
S90	GGGAGACAAGAATAAACGCTCAA	TGCTAACTTTAACAAACTCACCATTTC	TTCGACAGGAGGCTCACAACAGGC
S105	GGGAGACAAGAATAAACGCTCAA	GAAGTTGTAGCACCATTGGCACGG	TTCGACAGGAGGCTCACAACAGGC
S106	GGGAGACAAGAATAAACGCTCAA	TACACTTAAAAAGTAATTATGGGCG	TTCGACAGGAGGCTCACAACAGGC

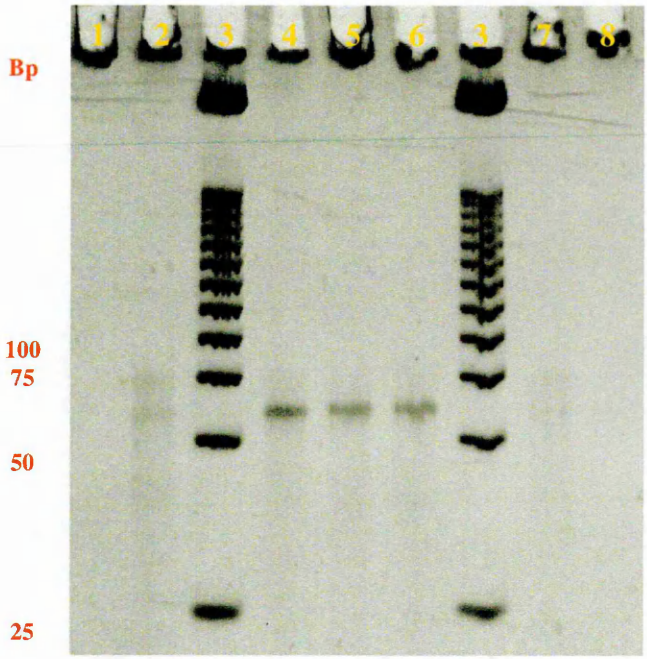
Appendix 1.0. Forty aptamer sequences from the SimpLex selection method. Sequences obtained from the 1.5 M NaCl concentration and 95°C samples for both MUC1/Y 10mer and 20mer peptides.

Sequences and predicted structures of the control aptamers obtained from the m-fold programme.

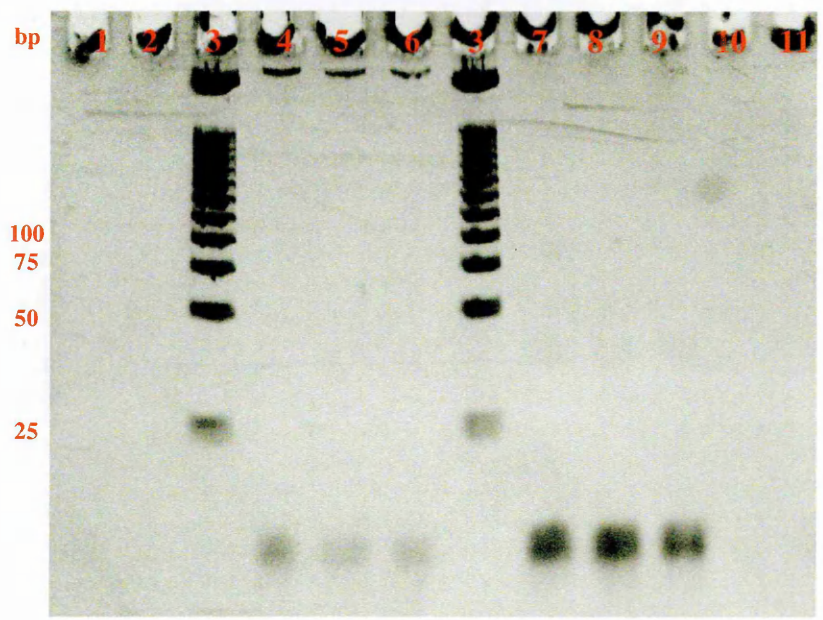


Appendix 2.0. The predicted secondary structures for the control aptamers. S11b-smd of 25nt (ATGCCAAATCGGCACGACTGAGAAT). S11c-smd of 26nt (CCGTATGCGGGCTATAATAGCACGAA). S75b-smd of 25nt (TAGTTAACGCCTAGGGTGTTTATAG). SP68 of 25nt (CGTCAGTATTTTGGGATGCCCGGGT). S2.2 (MUC1) of 25nt (GCAGTTGATCCTTTGGATACCCTGG).

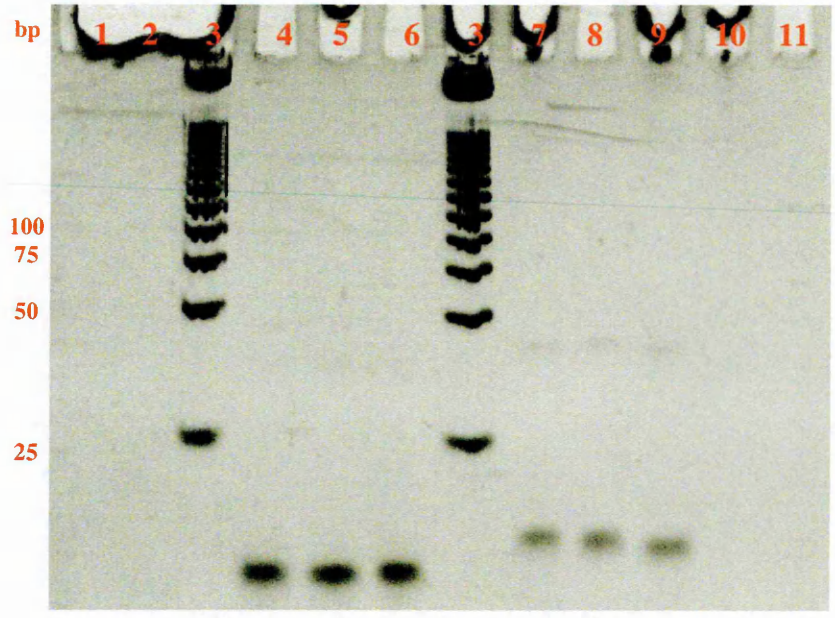
Analysis of aptamer-peptide mixtures by EMSA.



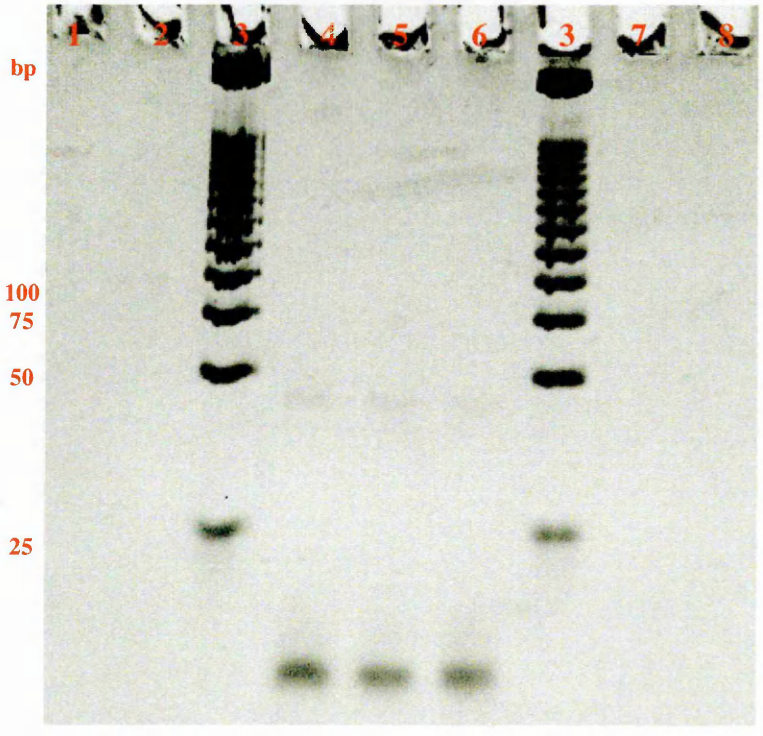
Appendix 2.1. Analysis of the S75a-peptide mixture. 14% native PAGE: (**1**) free 10 mer peptide (**2**) S75a + 10 mer at 95°C, (**3**) 25 bp ladder, (**4**) S75a + 10 mer, (**5**) free S75a, (**6**) S75a + 20 mer, (**7**) S75a + 20 mer at 95°C, (**8**) free 20mer peptide.



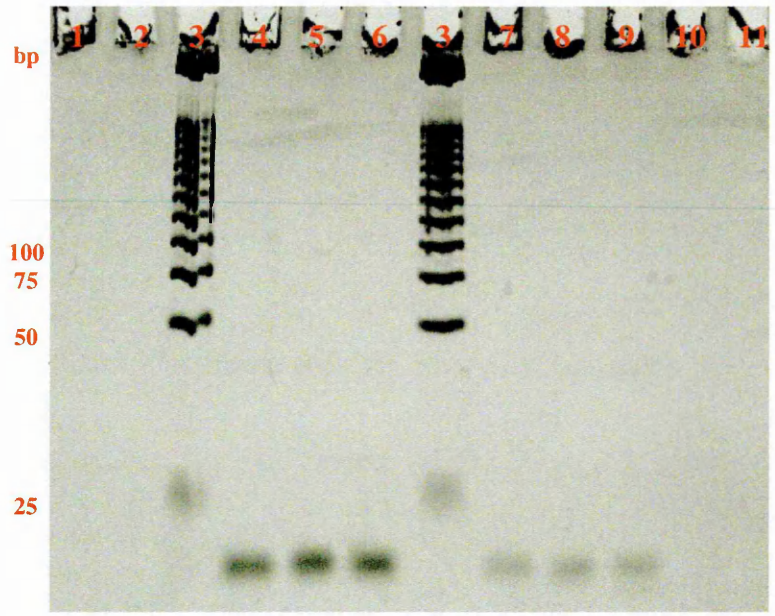
Appendix 2.2. Analysis of the S51b and S51c-peptide mixture. 18% native PAGE: S51b: (**1**) S51b + 10 mer at 95°C, (**2**) S51b + 20 mer at 95°C, (**3**) 25 bp ladder, (**4**) S51b + 10 mer, (**5**) free S51b, (**6**) S51b + 20 mer, **S51c:** (**7**) S51c + 10 mer, (**8**) free S51c, (**9**) S51c + 20 mer, (**10**) S51c + 10 mer at 95°C, (**11**) S51c + 20 mer at 95°C.



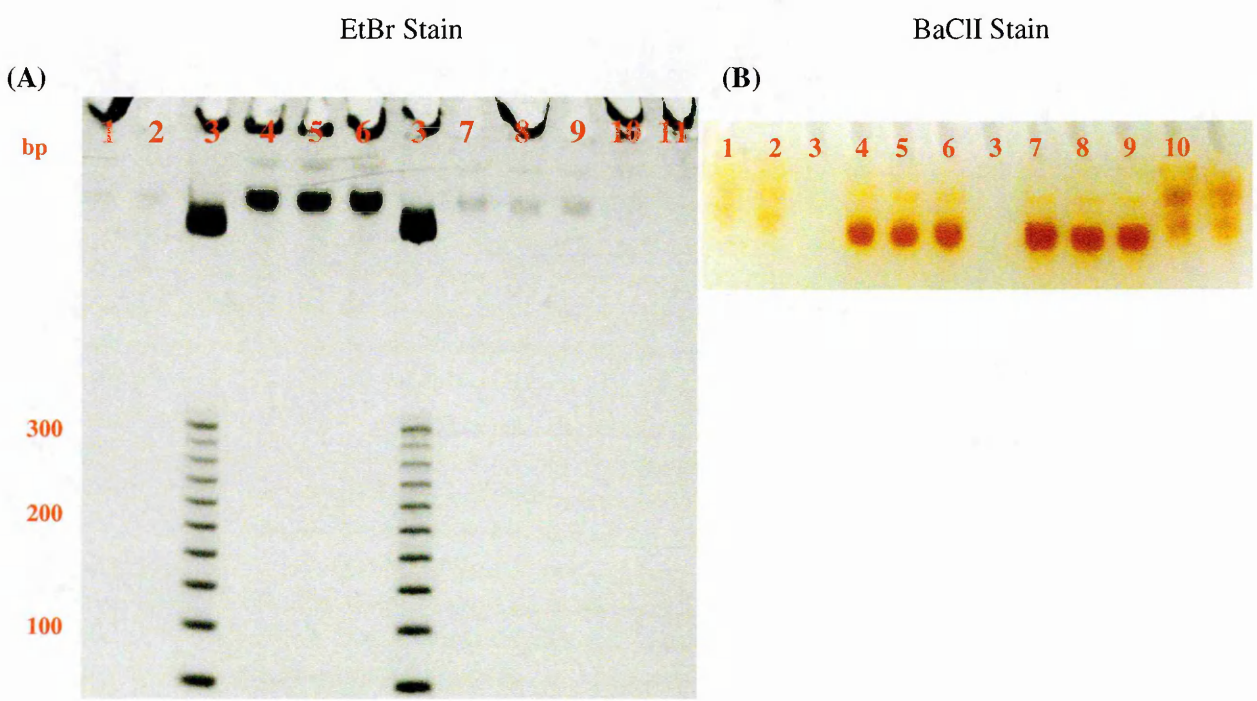
Appendix 2.3. Analysis of the S75b and S75b-smd-peptide mixture. 18% native PAGE: S75b: (1) S75b + 10 mer at 95°C, **(2)** S75b + 20 mer at 95°C, **(3)** 25 bp ladder, **(4)** S75b + 10 mer, **(5)** free S75b, **(6)** S75b + 20 mer, **S75b-smb: (7)** S75b-smd + 10 mer, **(8)** free S75b-smd, **(9)** S75b-smd + 20 mer, **(10)** S75b-smd + 10 mer at 95°C, **(11)** S75b-smd + 20 mer at 95°C. Smd = scrambled aptamer.



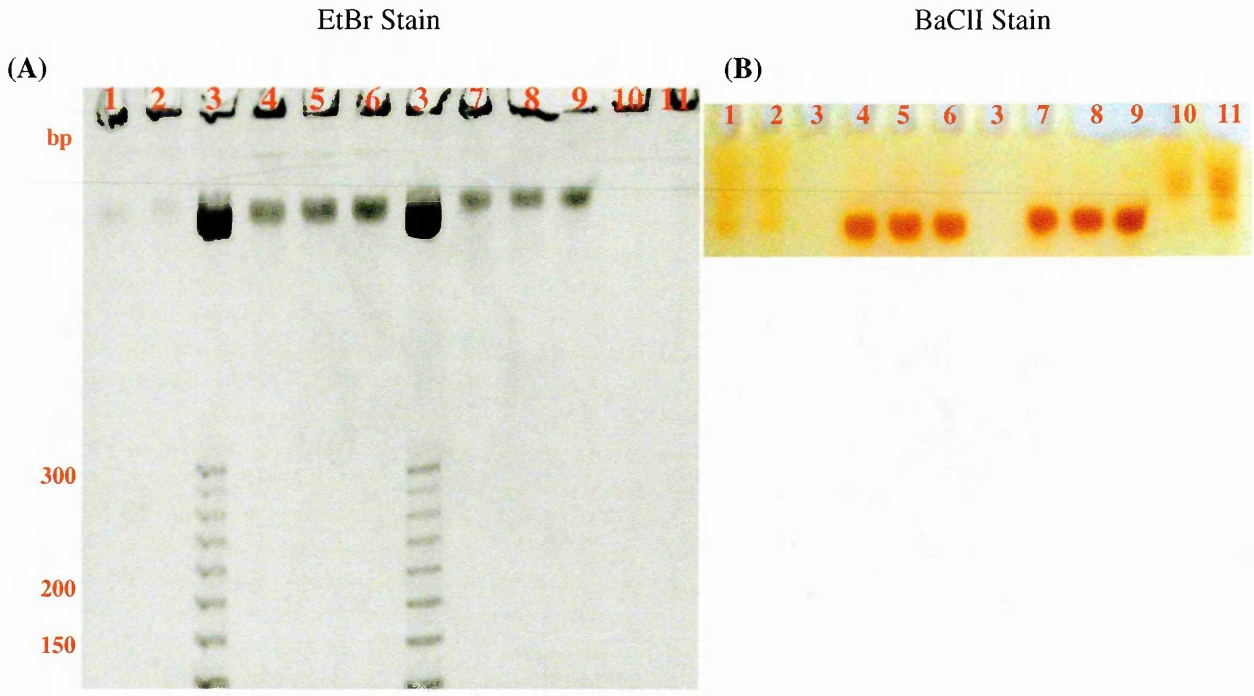
Appendix 2.4. Analysis of the S75c-peptide mixture. 18% native PAGE: (1) free 10 mer peptide **(2)** S75c + 10 mer at 95°C, **(3)** 25 bp ladder, **(4)** S75c + 10 mer, **(5)** free S75c, **(6)** S75c + 20 mer, **(7)** S75c + 20 mer at 95°C, **(8)** free 20mer peptide.



Appendix 2.5. Analysis of the SP68 and MUC1-peptide mixture. 18% native PAGE: SP68: (**1**) SP68 + 10 mer at 95°C, (**2**) SP68 + 20 mer at 95°C, (**3**) 25 bp ladder, (**4**) SP68 + 10 mer, (**5**) free SP68, (**6**) SP68 + 20 mer, **MUC1:** (**7**) MUC1 + 10 mer, (**8**) free MUC1, (**9**) MUC1 + 20 mer, (**10**) MUC1 + 10 mer at 95°C, (**11**) MUC1+ 20 mer at 95°C.



Appendix 2.6. Analysis of the pegylated S51b and S51c-peptide mixture. 7 % native PAGE: S51b-PEG: (**1**) S51b-PEG + 10 mer at 95°C, (**2**) S51b-PEG + 20 mer at 95°C, (**3**) 25 bp ladder, (**4**) S51b-PEG + 10 mer, (**5**) free S51b, (**6**) S51b-PEG + 20 mer, **S51c-PEG:** (**7**) S51c-PEG + 10 mer, (**8**) free S51c, (**9**) S51c-PEG + 20 mer, (**10**) S51-PEG c + 10 mer at 95°C, (**11**) S51c-PEG + 20 mer at 95°C.



Appendix 2.7. Analysis of the pegylated S75b and S75c-peptide mixture. 7 % native PAGE: S75b-PEG: (**1**) S75b-PEG + 10 mer at 95°C, (**2**) S75b-PEG + 20 mer at 95°C, (**3**) 25 bp ladder, (**4**) S75b-PEG + 10 mer, (**5**) free S75b, (**6**) S75b-PEG + 20 mer, **S75c-PEG:** (**7**) S75c-PEG + 10 mer, (**8**) free S75c, (**9**) S75c-PEG + 20 mer, (**10**) S75-PEG c + 10 mer at 95°C, (**11**) S75c-PEG + 20 mer at 95°C.

Analysis of aptamer-peptide interaction by affinity column chromatography.

Aptamer-Peptide	0.2M (%)	0.4M (%)	0.6M (%)	0.8M (%)	1.0M (%)	1.2M (%)	1.4M (%)	1.6M (%)	Bound aptamer (%)
S11a-10 mer	74.1 (± 1.3)	0	0	0	0	0	0	0	25.9 (± 1.3)
S11a-20 mer	70.7 (± 5.4)	0	0	0	0	0	0	0	29.3 (± 5.4)
S11b-10 mer	68.4 (± 1.7)	0	0	0	0	0	0	0	31.6 (± 1.7)
S11b-20 mer	78.5 (± 4.1)	0	0	0	0	0	0	0	21.5 (± 4.1)
S11b-smd- 10 mer	72.9	0	0	0	0	0	0	0	27.1
S11b-smd- 20 mer	80.0	0	0	0	0	0	0	0	20.0
S11b-PEG- 10 mer	79.5	0	0	0	0	0	0	0	20.5
S11b-PEG- 20 mer	52.6	0	0	0	0	0	0	0	47.5
S11c-10 mer	60.9 (± 3.5)	0	0	0	0	0	0	0	39.1 (± 3.5)
S11c-20 mer	64.6 (± 2.5)	0	0	0	0	0	0	0	35.4 (± 2.5)
S11c-smd- 10 mer	71.9 (± 4.5)	0	0	0	0	0	0	0	28.1 (± 4.5)
S11c-smd- 20 mer	75.0 (± 12.4)	0	0	0	0	0	0	0	25.0 (± 12.4)
S11c-PEG- 10 mer	22.3	0	0	0	0	0	0	0	77.7
S11c-PEG- 20 mer	16.8	0	0	0	0	0	0	0	83.2
S51a-10 mer	50.0 (± 23.1)	0	0	0	0	0	0	0	50.0 (± 23.1)
S51a-20 mer	54.8 (± 9.9)	0	0	0	0	0	0	0	45.2 (± 9.9)
S51b-10 mer	70.5 (± 6.2)	0	0	0	0	0	0	0	29.5 (± 6.2)
S51b-20 mer	75.5 (± 2.7)	0	0	0	0	0	0	0	24.5 (± 2.7)
S51b-PEG - 10 mer	20.2	0	0	0	0	0	0	0	79.8
S51b-PEG - 20 mer	40.2	0	0	0	0	0	0	0	59.8
S51c-10 mer	66.2 (± 11.6)	0	0	0	0	0	0	0	33.8 (± 11.6)
S51c-20 mer	74.3 (± 3.4)	0	0	0	0	0	0	0	25.7 (± 3.4)
S51c-PEG - 10 mer	10.7	0	0	0	0	0	0	0	89.3
S51c-PEG - 20 mer	23.8	0	0	0	0	0	0	0	76.2
SP68-10 mer	73.6 (± 11.8)	0	0	0	0	0	0	0	26.4 (± 11.8)

SP68-20 mer	74.6 (± 4.3)	0	0	0	0	0	0	0	25.4 (± 4.3)
S75a-10 mer	69.5 (± 3.1)	0	0	0	0	0	0	0	30.5 (± 3.1)
S75a-20 mer	68.6 (± 5.9)	0	0	0	0	0	0	0	31.4 (± 5.9)
S75b-10 mer	66.8 (± 5.5)	0	0	0	0	0	0	0	33.3 (± 5.5)
S75b-20 mer	75.3 (± 9.8)	0	0	0	0	0	0	0	24.7 (± 9.8)
S75b-smd - 10 mer	71.1 (± 10.8)	0	0	0	0	0	0	0	28.9 (± 10.8)
S75b-smd - 20 mer	78.4 (± 8.0)	0	0	0	0	0	0	0	21.6 (± 8.0)
S75b-PEG - 10 mer	59.0	0	0	0	0	0	0	0	41.0
S75b-PEG - 20 mer	62.8	0	0	0	0	0	0	0	37.2
S75c-10 mer	74.6 (± 6.8)	0	0	0	0	0	0	0	25.4 (± 6.8)
S75c-20 mer	71.2 (± 6.1)	0	0	0	0	0	0	0	28.8 (± 6.1)
S75c-PEG - 10 mer	45.6	0	0	0	0	0	0	0	54.4
S75c-PEG - 20 mer	44.8	0	0	0	0	0	0	0	55.2

Appendix 2.8. Aptamer binding with affinity chromatography. The percentages of each aptamer eluted and bound for both MUC1/Y peptides (10 mer and 20mer). Values are expressed as mean % (n =3), with the exception of the pegylated and S11b-smd (n = 1).

Analysis of aptamer-peptide mixtures by DNA thermal denaturation.

Aptamer	Thermal Denaturation Cycles		
	Cycle 1	**Cycle 2**	**Cycle 3**
S11a	35°C	35°C	35°C
S11a_10mer	43°C	23°C	13°C
S11a_20mer	30°C	12°C	< 5°C
S11b	21°C	21°C	21°C
S11b_10mer	7°C ; 20°C	10°C	7°C
S11b_20mer	16°C	< 3°C	< 3°C
S11b-smd	21°C	21°C	21°C
S11b-smd_10mer	20°C ; 35°C	17°C	12°C
S11b-smd_20mer	16°C	10°C	8°C
S11b-PEG	20°C ; 40°C	20°C ; 40°C	20°C ; 40°C
S11b-PEG_10mer	23°C	15°C	14°C
S11b-PEG_20mer	19°C	10°C	9°C
S11c	20°C ; 45°C	20°C ; 45°C	20°C ; 45°C
S11c_10mer	26°C	10°C; 22°C	15°C; 24°C
S11c_20mer	16°C	5°C	7°C

S11c-smd	47°C	47°C	47°C
S11c-smd_10mer	42°C	23°C	20°C
S11c-smd_20mer	36°C	13°C	< 3°C
S11c-PEG	20°C ; 51°C	20°C ; 51°C	20°C ; 51°C
S11c-PEG_10mer	35°C	20°C	12°C
S11c-PEG_20mer	19°C	13°C	5°C
S51a	30°C; 52°C	30°C; 52°C	30°C; 52°C
S51a_10mer	20°C; 36°C	17°C	15°C
S51a_20mer	29°C	<5°C	<5°C
S51b	18°C	18°C	18°C
S51b_10mer	25°C	7°C	10°C
S51b_20mer	21°C	12°C	<5°C
S51b-PEG	19°C	19°C	19°C
S51b-PEG_10mer	33°C	10°C	<5°C
S51b-PEG_20mer	23°C	24°C	21°C
S51c	18°C; 55°C	18°C; 55°C	18°C; 55°C
S51c_10mer	37°C	21°C	11°C
S51c_20mer	30°C	<5°C	<5°C
S51c-PEG	20°C; 61°C	20°C; 61°C	20°C; 61°C
S51c-PEG_10mer	36°C; 48°C	11°C	10°C
S51c-PEG_20mer	20°C; 50°C	15°C	<5°C
MUC1	18°C; 40°C	18°C; 40°C	18°C; 40°C
MUC1_10mer	29°C	16°C	12°C
MUC1_20mer	22°C	8°C	8°C
SP68	25°C	25°C	25°C
SP68_10mer	27°C	8°C	7°C
SP68_20mer	29°C	7°C	6°C
S75a	20°C; 35°C; 59°C	20°C; 35°C; 59°C	20°C; 35°C; 59°C
S75a_10mer	20°C; 43°C	8°C; 25°C	10°C
S75a_20mer	28°C	10°C; 25°C; 59°C	8°C
S75b	19°C	19°C	19°C
S75b_10mer	19°C	7°C	<3°C
S75b_20mer	23°C	<5°C	<5°C
S75b-smd	20°C; 47°C	20°C; 47°C	20°C; 47°C
S75b-smd_10mer	22°C	16°C	12°C; 27°C
S75b-smd_20mer	24°C	9°C	5°C; 22°C
S75b-PEG	20°C; 40°C	20°C; 40°C	20°C; 40°C
S75b-PEG_10mer	19°C; 34°C	15°C	10°C; 28°C
S75b-PEG_20mer	28°C	20°C	<5°C
S75c	22°C; 43°C; 63°C	22°C; 43°C; 63°C	22°C; 43°C; 63°C
S75c_10mer	20°C; 30°C	14°C; 24°C	10°C; 25°C; 37°C
S75c_20mer	22°C	13°C	<3°C
S75c-PEG	21°C; 35°C; 63°C	21°C; 35°C; 63°C	21°C; 35°C; 63°C
S75c-PEG_10mer	39°C	19°C	10°C; 22°C
S75c-PEG_20mer	22°C	16°C	<5°C

Appendix 2.9. Melting temperatures of the MUC1/Y aptamers. Melting temperatures of each aptamer for all three TDC with both MUC1/Y peptides (10 mer and 20mer).