# Kent Academic Repository

**Paul, Alexandra Rosie (2022)** *A Novel Strategy for the Synthetic Selection of Enhanced Therapeutic Aptamers.* **Doctor of Philosophy (PhD) thesis, University of Kent,.**

# A Novel Strategy for the Synthetic Selection of Enhanced Therapeutic Aptamers

By

## Alexandra Rosie Paul

Division of Natural Sciences, University of Kent

*A thesis submitted for the degree of Doctor of Philosophy*

**June 2022**

# Abstract

The epidermal growth factor receptor protein (EGFR) promotes tumorigenesis in many cancers, including breast, lung, colon and glioblastoma.[1] Deregulation of EGFR activity in cancer is mainly due to point mutation, kinase domain deletion or gene amplification.[2] It is essential to have EGFR-targeted therapies because it is among the most commonly altered genes in cancer.[3] Aptamers are single-stranded oligonucleotide sequences that bind with high affinity and specificity to diverse targets.[4] The recognition capacity of aptamers can be harnessed for therapeutic agents. The chemistry of aptamers is largely limited to that of nucleic acids, and although non-natural modifications of nucleic acids are known to enhance aptamer affinity,[5] there is not yet a technology for selecting the right modifications amongst billions of possibilities.

This project aims to develop the first general method for the discovery of nucleoside modifications that increase aptamer binding efficacy. A library will be created of over 65,000 different chemical modifications on a known aptamer sequence (MinE07),[6] which binds to EGFR protein. The modifications will use different types of chemistry to see how they will affect the binding and folding of the aptamer MinE07, aiming to improve the binding between MinE07 and EGFR. These will be attached on beads such that each bead displays millions of copies of a single aptamer sequence. The selection will then be carried out using a flow cytometer (FACS) to separate out the aptamer sequences with the highest affinity for EGFR from the one-bead-one-sequence aptamer library. These will then be identified using mass spectrometry. These top selected aptamer sequences will be subjected to biophysical and biological testing. This research provides a new method for synthesising and screening in a short space of time large drug candidate libraries producing new drug candidates and impacting drug discovery processes.

# Declaration

I declare that all the work presented in this thesis submission was my own work and written in my own words. It has not been submitted to any other institution for a qualification or for any other degree.

Alexandra Rosie Paul

June 2022

# COVID-19 Mitigation

Covid-19 pandemic has impacted my research in a number of ways. It has restricted my access to both facilities and materials and thus limited my ability to generate and collect data and thus haulting my research progression. Specifically, From March 2020 – April 2021 we had limited or no access to the labs and equipment, totalling to over 7 months of no access. Furthermore, there was disruption of supply chains, which then limited my ability to undertake specific parts of my project in a timely manner. Together, these issues massively impacted my research as this project required practical experiments in a lab full time with specialist equipment across the department.

Dr Christopher J. Serpell

10.06.2022

Prof Michelle D. Garrett

09.06.2022

# Acknowledgements

*Dedicated to Rosalind, Matthew and Bethany Paul, this achievement is all yours!*

Firstly, I would like to thank my supervisors Dr Christopher Serpell and Prof Michelle Garrett for giving me this opportunity to work in your labs. Without your immense support, mentorship, and kindness I would not have made it to the end of my PhD. Your love of science and passion for research has been very inspirational and has led me to want to stay on this path. I owe the valuable skills and knowledge that I have acquired over the past few years to you both.

I wish to thank the University of Kent Vice-Chancellor Scholarship scheme for funding my research. Thank you to the Royal Society of Chemistry for awarding me with the Research Enablement Grant, allowed for an extra 6 months in the lab which has been invaluable in the progress of this project. Thank you to Centauri Therapeutics for providing me with consumable money for the first year of my project. I am very grateful for the mentorship and teaching from Helen Lavender during my time with Centauri Therapeutics.

Further thanks go to Edith Blackburn for all her help with flow cytometry, and mentorship in the biology labs. Kevin Howland for all his time and teaching on the mass spectrometer and Ian Brown for all his help with microscopy.

Thank you to everyone in the Serpell/Garrett labs for everything you have taught me and supported me with. I especially thank my work wife Bini Claringbold and my fume hood partner Emerald Taylor for the huge amount of love and support during my PhD.

Huge thank you to my best friends who are forever cheering me on, I love you. Thank you to my incredibly supportive, loving, and sacrificial parents. This achievement is all yours. And a final thank you to l'amore della mia vita, Mario Falsaperna. La vita è migliore con te.

# Dissemination

## Grants

Royal Society of Chemistry, Research Enablement Grant – E21-9810231002, £9980.00

'A First Strategy for the Synthetic Selection of Enhanced Therapeutic Aptamers', Co-PI, (Sept 2021- Sept 2022)

## Publications in Preparation

A novel strategy for the synthetic selection of enhanced therapeutic aptamers. **Paul, A. R.**, Lavender, H., Garrett, M. D., Serpell, C. J.

Understanding sequence-programmed self-assembly in oligophophoesters with coarse-grained and atomistic modelling. Piskorz, T, **Paul, A. R**., Claringbold, B. R., Garrett, M. D., Duarte, F., Serpell, C. J.

## Oral Presentations

"A Novel Strategy for the Synthetic Selection of Enhanced Therapeutic Aptamers." Bring Chemical Biology to Cancer Research, Royal Society of Chemistry & British Association of Cancer Research, Virtual (Dec 2021)

## Flash Talks

"Synthetic Selection of Enhanced Therapeutic Aptamers." 7th International INSOAP Symposium on Aptamers, Virtual (Sept 2020)

# Poster Presentations

"The First Strategy for the Synthetic Selection of Enhanced Therapeutic Aptamers." Biological and Medicinal Chemistry, BMCS Postgraduate Symposium XV, Virtual (Dec 2021)

"Synthetic Selection of Enhanced Therapeutic Aptamers." 7th International INSOAP Symposium on Aptamers, Virtual (Sept 2020)

"Synthetic Selection of Enhanced Therapeutic Aptamers." RSC Macrocyclic and Supramolecular Chemistry Meeting, University of Kent, Canterbury (Dec 2019)

"Synthetic Selection of Enhanced Therapeutic Aptamers." London Polymer Group Symposium, University of Kent, Canterbury (April 2019)

# Table of Content

# List of Abbreviations

| | |
|---|---|
| **AHR** | Airway Hyper Responsiveness |
| **APC** | Allophycocyanin |
| **AR** | Amphiregulin |
| **BTC** | Betacellulin |
| **c-Cbl** | Casitas B-lineage Lymphoma proto-oncogene |
| **c-Fos** | FBJ murine osteosarcoma viral oncogene homolog |
| **c-Myc** | Myelocytomatosis viral oncogene homolog |
| **c-Src** | Sarcoma viral oncogene homolog |
| **CCR5** | CC-Chemokine Receptor 5 |
| **CDC42** | Cell Division Control protein 42 homolog |
| **Cdk2** | Cyclin-dependent kinase 2 |
| **CE-SELEX** | Capillary Electrophoresis-SELEX |
| **Cy5.5** | Cyanine5.5 NHS ester |
| **DAG** | Diacylglycerol |
| **DNA** | Deoxyribonucleic Acid |
| **DCM** | Dichloromethane |
| **DMF** | Dimethylformamide |
| **DMSO** | Dimethyl Sulfoxide |
| **DMTrCl** | 4,4'-Dimethoxytriyl Chloride |
| **DSC** | Differential Scanning Calorimetric |
| **EGF** | Epidermal Growth Factor |
| **EMT** | Epithelial-to-Mesenchymal Transition |

| | |
|---|---|
| **EPG** | Epigen |
| **EPR** | Epiregulin |
| **ErbB1/HER-1** | ErbB1/Human Epidermal Growth Factor Receptor |
| **ErbB2/HER-2** | ErbB2/ Human Epidermal Growth Factor Receptor |
| **ErbB3/HER-3** | ErbB3/ Human Epidermal Growth Factor Receptor |
| **ErbB4/HER-4** | ErbB4/ Human Epidermal Growth Factor Receptor |
| **EGFR** | Epidermal Growth Factor Receptor |
| **FACS** | Fluorescent Activated Cell Sorter |
| **FDA** | Food and Drug Administration |
| **FITC** | Fluorescein Isothiocyanate |
| **FRET** | Fluorescence Resonance Energy Transfer |
| **FSC** | Forward Scatter |
| **GFP** | Green Fluorescent Protein |
| **GP43** | Bacteriophage T4 DNA Polymerase |
| **GRB2** | Growth Factor Receptor-bound Protein |
| **GSK3 beta** | Glycogen Synthase Kinase 3 beta |
| **GST** | Glutathione S-Transferase |
| **HB-EGF** | Heparin Binding EGF-like Growth Factor |
| **HCV** | Hepatitis C Virus |
| **HER1/2/3/4** | Human Epidermal Growth Factor Receptor |
| **HFNAPs** | Highly Functionalized Nucleic Acid Polymers |
| **hFc1** | Fc fragment of human IgG1 |
| **HIFA** | 2-Hydroxyisophthalate Formaldehyde Acetal |
| **HIV** | Human Immunodeficiency Virus |
| **HTS** | High Throughput Screening |

| | |
|---|---|
| **ILK** | Integrin-Linked kinase |
| **IKK** | Nuclear factor NF-kappa-B Inhibitor Kinase |
| **ITC** | Isothermal Titration Calorimetry |
| **JAK1** | Janus Kinase 1 |
| **JAK2** | Janus Kinase 2 |
| **JNK1** | Mitogen-activated protein kinase 8 |
| **JNK2** | Mitogen-activated protein kinase 9 |
| $K_a$ | Affinity Constant |
| **kDa** | Kilodaltons |
| **KCN** | Potassium Cyanide |
| **LNA** | Locked Nucleic Acids |
| **LOD** | Limit of Detection |
| **MAP** | Mitogen Activated Protein |
| **MEKK1** | Mitogen-activated Protein Kinase Kinase 1 |
| **MEKK2** | Mitogen-activated Protein Kinase Kinase 2 |
| $MgSO_4$ | Magnesium Sulfate |
| **MKP-1** | Map Kinase Phosphatase -1 (DUSP1) |
| **MKP-2** | Map Kinase Phosphatase - 2 (DUSP 4) |
| **MLK2** | Mitogen-Activated Protein Kinase Kinase Kinase 10 |
| **MMP-2/9** | Matrix metallopeptidases |
| **MS** | Mass Spectrometry |
| **mTBA** | Modified Thrombin Binding Aptamer |
| $N_2$ | Nitrogen |
| **Na** | Sodium |
| **NaCl** | Sodium Chloride |

| | |
|---|---|
| **NCK** | NCK adaptor protein 1 |
| **NF-kB** | Nuclear Factor kappa-B |
| **NGS** | Next Generation Sequencing |
| **NH$_3$/MeOH** | Ammonia in Methanol |
| **NRG** | Neuregulins |
| **NMR** | Nuclear Magnetic Resonance |
| **OBOC** | One Bead One Compound |
| **PAK1** | p21-Actviated Kinase 1 |
| **PCR** | Polymerase Chain Reaction |
| **PCSK9** | Proprotein Convertase Subtilisin/Kexin Type 9 |
| **PDGF-BB** | Platelet-Derived Growth Factor BB |
| **PerCP** | Peridinin-Chlorophyll-Protein |
| **PI3K** | Phosphoinositide 3-kinase |
| **PI3K reg class IA (p85)** | p85 subunit of phosphatidylinositol 3 kinase |
| **PKC** | Protein Kinase C |
| **PLCγ1** | Phospholipase C gamma 1 |
| **PtdIns(3,4,5)P3** | Phosphatidylinositol 3,4,5-triphosphate |
| **(PtdIns(4,5)P2)** | Phosphatidylinositol 4,5-biphosphate |
| **QCM** | Quartz Crystal Microbalance |
| **Rac1** | Ras-related C3 botulinum toxic substrate 1 |
| **Raf-1** | v-Raf-1 murine leukemia viral oncogene homolog 1 |
| **H-Ras** | GTPase v-Ha-ras Harvey rat sarcoma viral oncogene homolog |
| **RhoA** | Ras homolog gene family, member A |
| **RNA** | Ribonucleic Acid |
| **RNP** | Ribnonuleoprotein |

| | |
|---|---|
| **RT** | Rention Time |
| **SAR** | Structure Activity Relationship |
| **SELEX** | Systematic Evolution of Ligands by EXponential enrichment |
| **SH2** | Src Homology |
| **SHC** | SHC transforming protein 1 |
| **SOMAmers** | Slow Off-rate Modified Aptamers |
| **SOS** | Son Of Sevenlees homolog 1 |
| **SSC** | Side Scatter |
| **STAT3** | Signal Transducer and Activator of Transcription 3 |
| **TentaGel Microspheres** | TentaGel® M $NH_2$ Monosized Amino TentaGel Microspheres |
| **TGF-alpha** | Transforming Growth Factor alpha |
| **TGF-beta** | Transforming Growth Factor beta |
| **UV** | Ultraviolet |
| **VIS** | Visible |

# Chapter 1

## Aptamers, Combinatorial Chemistry, Flow Cytometry, Cancer, and the Epidermal Growth Factor Receptor

# Chapter 1

## 1.1. Aptamers

### 1.1.1. Aptamers

Aptamers are single-stranded oligonucleotides that mimic antibodies by folding into complex 3D shapes that bind with high affinity and specificity to diverse targets, non-covalently (figure 1.1).[4] Oligonucleotides are short nucleic acid molecules (15-60 nucleotides long) which are often generated by solid-phase phosphoramidite chemistry.[7] There are three main types of aptamers, ribonucleic acid (RNA), deoxyribonucleic acid (DNA) and peptide which all have similar properties. They can be manufactured as single-stranded molecules with any chosen sequence.[8]



Single Stranded Oligonucleotide          Folder Aptamer
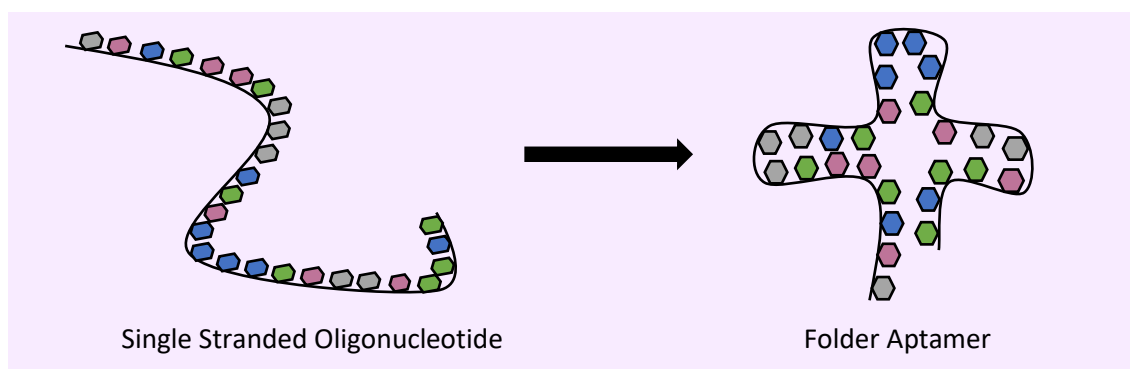
Figure 1.1. Aptamer in Single Stranded and Folded Conformations

So far, aptamers have been selected against viruses,[9] proteins,[10] polysaccharides,[11] bacteria,[12] toxins,[13] peptides,[14] small molecules,[15] amino acids[16] and whole cells.[17] Aptamers can discriminate their selected target with high precision; they have high binding affinity, high specificity and binding to the target is non-covalent, but strong. This recognition capacity of aptamers can be harnessed for therapeutic agents. Aptamers have a large range of potential targets compared with other molecules such as antibodies because they structurally conform to bind well.[18] Aptamers can be chemically modified to introduce the site-specific addition of

non-nucleotide linkers, fluorescent labels and different functionalities that help with conjugation. Non-natural modifications of nucleic acids are known to enhance aptamer affinity, however there is not a technology for selecting the right modifications amongst millions of possibilities.

### 1.1.1.1. DNA and RNA Aptamers

The most common types of aptamers are RNA and DNA-based. They are selected from a large pool of ~$10^{18}$ nucleic acid sequences. Each stage of the aptamer selection process involves binding, elution and finally amplification to identify selective, high-affinity aptamers that binds to a specific target. For DNA aptamer selection the starting library consists of single-stranded DNA oligonucleotides. RNA aptamer selection starts with a double stranded DNA library that is then transcribed to create the starting RNA library. For RNA aptamers, reverse transcription must be performed on each round of selection to facilitate polymerase chain reaction (PCR) of DNA and subsequent transcription for the next round of selection.[19] This means that DNA aptamers are easier and cheaper to make and are more stable products than RNA aptamers.[18] RNA is less stable than DNA because it has a 2'-hydroxyl group on the pentose ring which is very susceptible to hydrolysis (figure 1.2). RNA aptamers however can form more diverse and complex 3D structures, which allow them to have a greater number of conformations compared with DNA aptamers.[20]
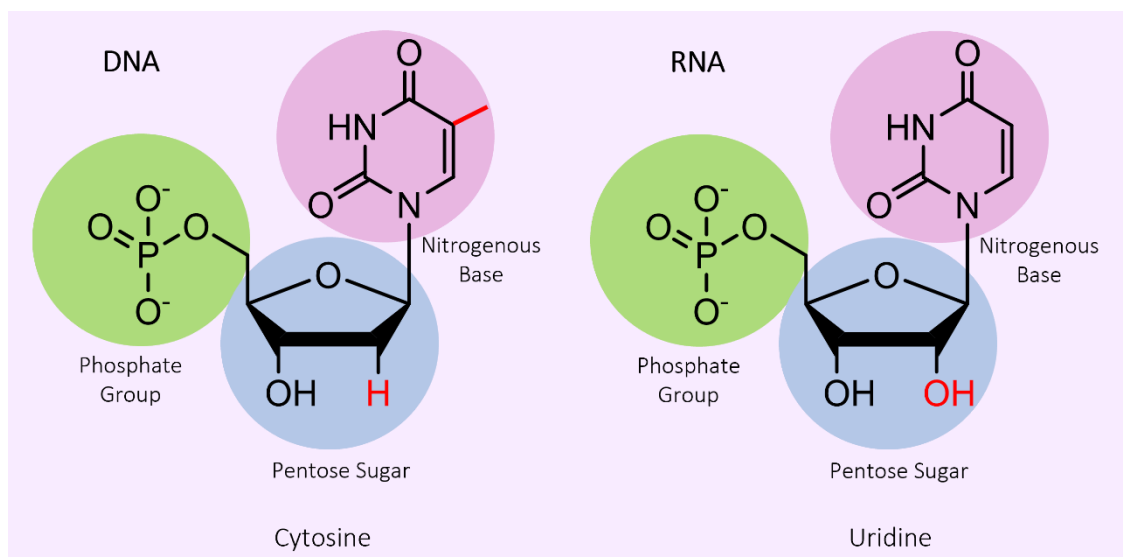
Figure 1.2. DNA nucleotide Vs RNA nucleotide Structure

### 1.1.1.2. Peptide Aptamers

Peptide aptamers were first developed in 1996 and were yet another alternative to antibodies,[21] and a concept originally introduced by Roger Brent. Peptide aptamers are small peptide molecules, 5-20 amino acids in length,[21] with a loop region which is variable, and a protein attached to both ends (figure 1.3). This loop region is where the peptide aptamer interacts to its target. The flexibility of peptide aptamers is low because of the target being joined to the loop region. This means that this aptamer is less effective than the DNA and RNA aptamers. Peptide aptamers do however have very high specificity. Peptide aptamers can be used to inhibit protein-protein interactions.[18] The binding affinity of constrained aptamers can be as much as 1000 times higher than the free peptide, attributed to the lower conformational entropy of the restricted peptide loop.[22] Peptide aptamers can be produced and selected *in vivo* through yeast two-hybrid and similar techniques. This makes them ideal candidates for interrogating intracellular targets in a physiological environment.[23]
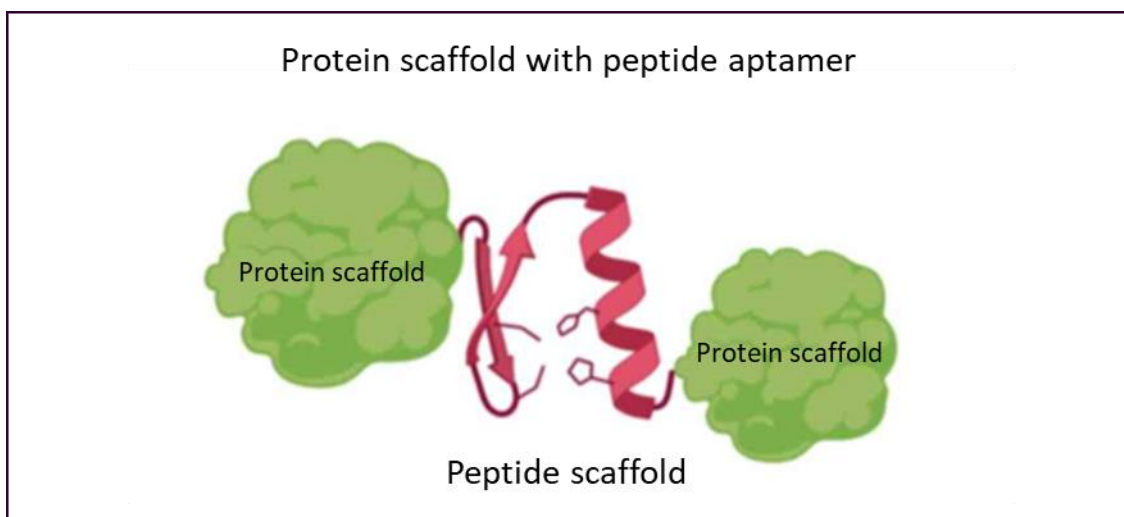
Figure 1.3. Protein scaffold with peptide aptamer.[24]

Table 1.1. Differences between DNA, RNA and Peptide aptamers.

|  | DNA Aptamers | RNA Aptamers | Peptide Aptamers |
|---|---|---|---|
| Complexity of structures | Form complex secondary and tertiary structures | Form complex secondary and tertiary structures | Structure constrained by peptide and protein scaffold |
| Diversity of 3D structures | Less diverse 3D structures than RNA aptamer | Form diverse 3D structures | 3D structure constrained by protein scaffold |
| Binding capacity | Bind to target with the entire sequence | Bind to target with the entire sequence | Bind to target via variable region only |
| Uses | Biosensor, Diagnostic, Therapeutic applications | Biosensor, Diagnostic, Therapeutic applications | Biosensor, Diagnostic, Therapeutic applications |

## 1.1.2. Aptamers Binding with their Targets

Nuclear magnetic resonance (NMR) spectroscopy and X-Ray crystallography have been used to analyse several aptamers. The results revealed that specificity and high affinity are achieved by the 3D structures of aptamers. Aptamers can be a variety of different structures, such as hairpins, bulges, stems and pseudoknots (figure 1.4).
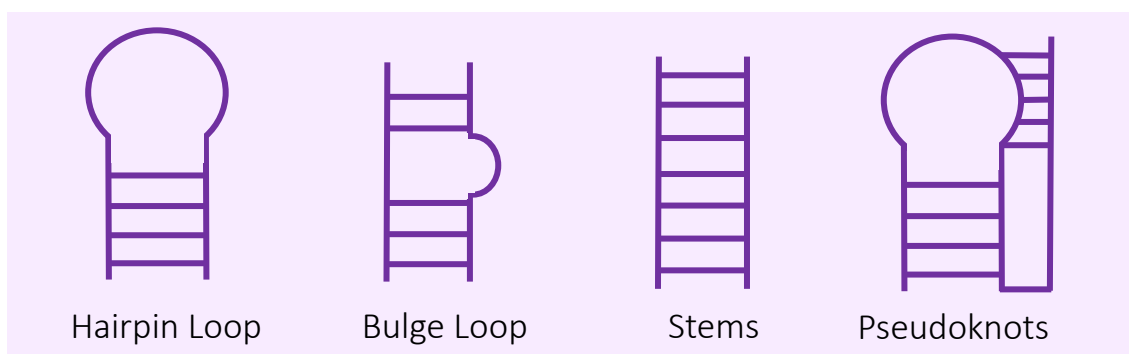


Figure 1.4. Hairpin loop, bulge loop structures[25], stems and Pseudoknot structures.[26]

### 1.1.2.1. Binding forces that facilitate the aptamer–target interactions

Hydrogen bonding, electrostatic interactions, hydrophobic interactions, π–π stacking and Van der Waals forces are the primary binding forces between aptamers and their targets.[27] The binding affinity may vary between each aptamer and its target and many different binding forces may be involved in one aptamer-target interaction. Conditions that may affect this interaction include pH, temperature, metal ion strength amongst others.

#### 1.1.2.1.1. Hydrogen bonding

Hydrogen bonding is the most common binding force between aptamers and their target molecules (figure 1.5).[28] RNA-to-protein interactions mostly comprise the side chains of proteins, which act principally as hydrogen bond donors, and the negatively charged phosphate groups of the backbone of the RNA, which act as an H-bond acceptors.[29] Nagatoishi *et al.* studied the thermal stability of DNA aptamers bound to thrombin by isothermal titration calorimetry

(ITC) to demonstrate the dominant role of hydrogen bonding in the aptamer-target binding process. The results confirmed that direct interactions with thrombin via hydrogen bonding are the main reason for the stability of the DNA aptamer G-quadruplex structure.[30] In 2000, Cowan *et al.* applied ITC to compare the contributions that electrostatic interactions and hydrogen bonding make to the binding and recognition of neomycin B molecules to the RNA aptamer. The results show that the primary mechanism is hydrogen bonding. At physiological ionic strength, around 80% of the binding energy originates from the hydrogen bonding in the A and B rings of the neomycin B molecules.[31] As well as facilitating the binding process, the inner hydrogen bonding of the aptamer is essential to the stability of the aptamer.[32] In 2008, Pagano *et al.* discovered that in the inner part of the modified thrombin binding aptamer (mTBA), T7 base forms intramolecular hydrogen bonds with the guanine bases of tetrads. However, there are none in the inner part of the non-modified TBA. mTBA has been shown to be more stable than TBA by differential scanning calorimetric (DSC).[33]
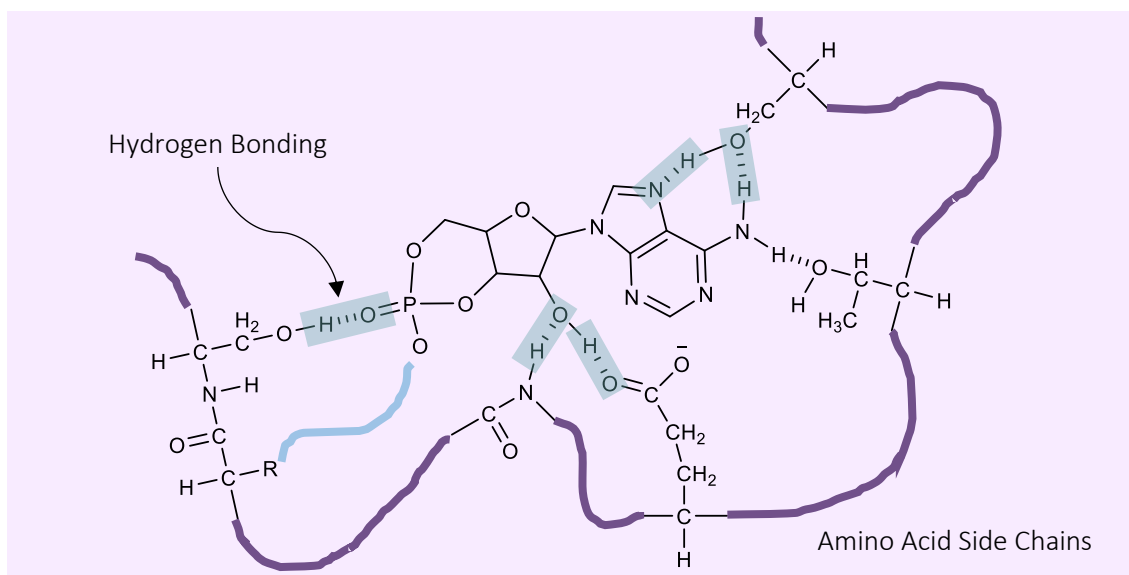


Figure 1.5. Hydrogen bonding between Nucleic Acids and Amino Acid Side Chains.

### 1.1.2.1.2. Electrostatic Interactions

Electrostatic interaction is another force that facilitates the binding between aptamer and target. The electrostatic interactions between nucleic acids and proteins occur due to the negatively charged phosphate backbones of the aptamer and the positively charged amino acids of the protein (figure 1.5).[34] Some crystal structures have shown that RNA aptamers bind to target proteins mainly via electrostatic forces.[35,36] The electrostatic force is considered to result in non-specific binding because of the lack of spatial directionality.[37] In 2012 Rangnekar *et al.* found that they were able to attach an ssDNA tail to a DNA aptamer and the aptamer activity would increase compared to when the aptamer was free in solution.[38] They concluded that this was because of the electrostatic interactions between the negatively charged tail and the positively charged amino acids on the surface of thrombin. The salt concentration in a reaction can commonly affect the intensity of the electrostatic interactions. In 2011 Lin *et al.* discovered a decrease in the affinity constant ($K_a$) of a 15-mer DNA aptamer binding to its thrombin target when more $Na^+$ ions were added to the solution. They concluded that the 15-mer aptamer bound to the fibrinogen binding site of thrombin using a G-quadruplex structure primarily through electrostatic interactions.[39]

### 1.1.2.1.3. Hydrophobic Phenomena

Hydrophobic effects can contribute to the binding between aptamers and their targets and also stabilise the aptamer 3D structure. In 2013 Chou *et al.* found that DNA sequences rich with guanine form stable G-quadruplex structures via G-tetrad hydrogen-bonding and hydrophobic stacking.[40] Lin *et al.* stated that a 29-mer aptamer bound to the heparin binding site of thrombin using a duplex structure and was driven primarily due through hydrophobic effects.[39] Slow off-rate modified aptamers (SOMAmers) bind to targets by complementary shapes and hydrophobic modifications at the binding interface. This is mostly done by hydrophobic aromatic rings which mainly interact with the aliphatic side chains of the amino acid residues of

the protein target as shown in the interaction between the PDGF-BB dimer complex and two SOMAmers (figure 1.6).[41]

In 2017 Ren *et al.* discovered the functional interactions of cytokine IL-1α with modified DNA aptamer SL1067 by solving the high-resolution structure of the binding complex.[42] They established that the interface of the binding complex consists of a network of hydrophobic interactions. This alluded that the majority of the binding energy came from a cluster of hydrophobic moieties on the surface of DNA aptamer SL1067 and the hydrophobic regions of amino acids on the IL-1α surface. It was also found that hydrogen bonding was involved in the binding.
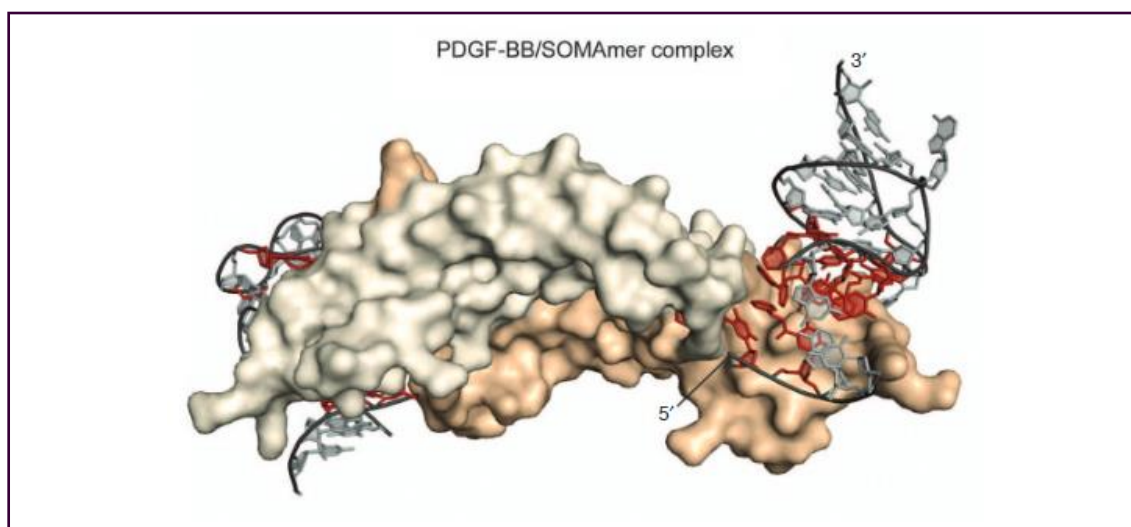


Figure 1.6.  Crystal structure of SOMAmers bound to the PDGF-BB dimer (platelet-derived growth factor B).[41]

### 1.1.2.1.4 π–π Stacking

π–π stacking is the stabilising non-covalent interaction commonly observed between parallel aromatic rings. π–π stacking interactions transpire in a large variety of chemistry and are known for influencing self-assembly processes, inserting molecules into DNA, involvement in host-guest complexes and induction of side chain interactions in proteins.[43] In 2011 Bernard *et al.* reported that the malachite green aptamer (MGA) recognises and binds to its target, malachite green molecule, mainly via π–π stacking and electrostatic interactions.[44] In the binding complex,

a base quadruplet and a Watson-Crick base pair act as a stacking platform for malachite green. The other residues next to the binding pocket act as anchors and linkers to enable the binding.

### 1.1.2.1.5. Van der Waals Forces

Van der Waals forces consist of distance-dependent attractions and repulsions between atoms or molecules. They are non-directional noncovalent forces and are generally weaker than other bonding forces such as hydrogen bonding and electrostatic interactions. It has been found that Van der Waals forces are involved in aptamer to target binding. In 2014, Zhang *et al.* discovered that both hydrogen bonding and Van der Waals forces were at work in the binding of organophosphorus pesticides to their aptamer.[45] In 2010, Normura *et al.* investigated the structural interactions between the Fc fragment of human IgG1 (hFc1) bound to an anti-Fc RNA aptamer using a 2.15 Å crystal structure of the complex. It was found that the binding between hFc1 and the RNA aptamer was facilitated primarily by multiple weaker interactions, such as Van der Waals forces and hydrogen bonding, instead of electrostatic forces.[46]

### 1.1.2.2. Aptamer Interactions with Proteins
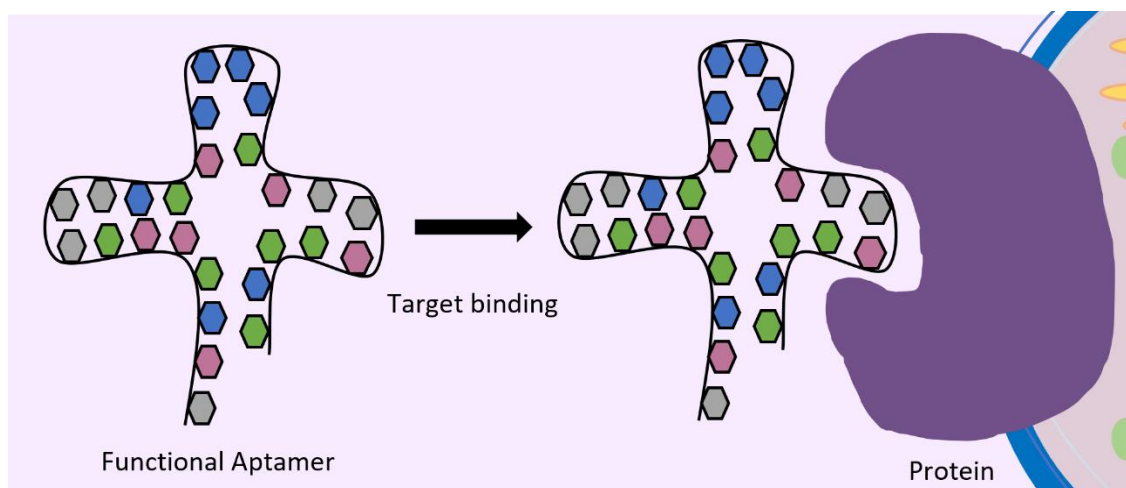


Figure 1.7. Aptamers binding to Protein Targets.[47]

RNA and DNA aptamers bind well with proteins because the phosphate back-bone of the aptamer has conformational flexibility, which allows shape complementarity with the target protein interface and so most aptamers bind to their targeted protein without any substantial

change in their conformation (figure 1.7). Aptamers have greater affinities for their protein targets than antibodies have for their antigen.[48] It is the high affinity of aptamers for their target that allows them to be used in applications such as protein purification, protein profiling clips, recognising and targeting the chosen protein from a mixture of them.[49]

## 1.1.3. Aptamer Stability

DNA aptamers are more stable than RNA aptamers. In a controlled environment outside of a living organism the half-life of an RNA aptamer in plasma is a few seconds compared to DNA aptamers, which is approximately 30-60 minutes.[50] RNA is a short term messenger in the cell and is chemically unstable because of its 2'-hydroxyl group. The 2'-hydroxyl group makes RNA susceptible to hydrolysis (figure 1.7).[19] The 2'-hydroxyl group also allows for the catalysis of RNA strand scission by endoribonuclease.[51] Because of this, most RNA aptamers are chemically modified by either the design of the library or during the selection amplification, which helps enhance RNA aptamer stability.
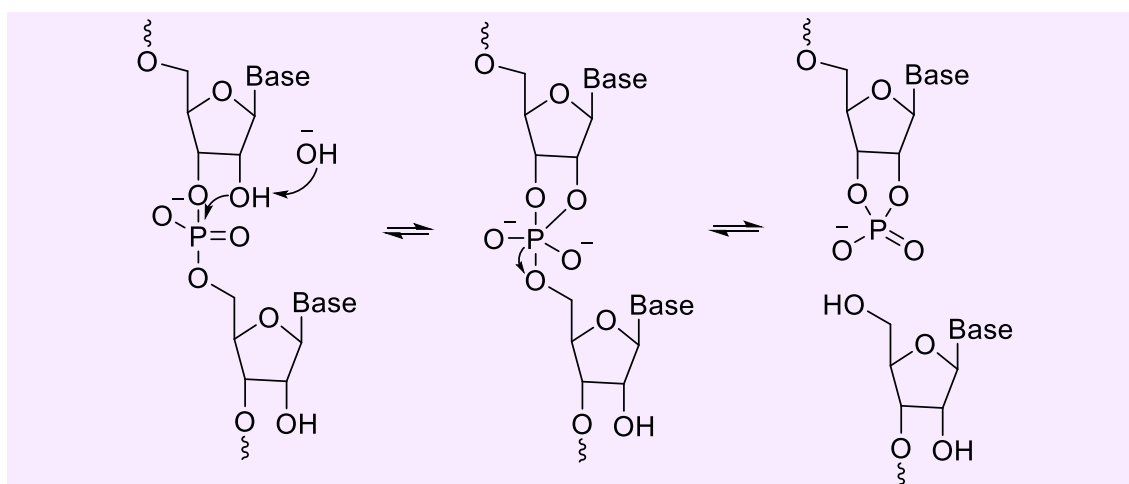


Figure 1.8. Hydrolysis of RNA

Protecting groups can be used on the 2'-hydroxyl position, an example is 2-hydroxyisophthalate formaldehyde acetal (HIFA). This protecting group works since the bis ester is stable in acidic

conditions that are used for the removal of dimethoxytrityl groups during chain elongation. Then in the final deprotection step, which is in alkali conditions, there is chain cleavage from the support, which removes the base and phosphate protecting groups. This compound is then converted in to the bis carboxylate which can be removed rapidly when treated with mild acid.[52] Another example is O-tert-butyldimethylsilyl chloride as shown in figure 1.9.  After such modification, RNA aptamers are either equal or better than DNA aptamers in terms of stability.[53] The most common modifications used for the therapeutic aptamers are the 2'-Fluoro and the 2'-O-Methyl. The 2'-Fluoro modifications enhance base pairing stability and produce aptamers with better affinities as opposed to the 2'-amino aptamers.[54]
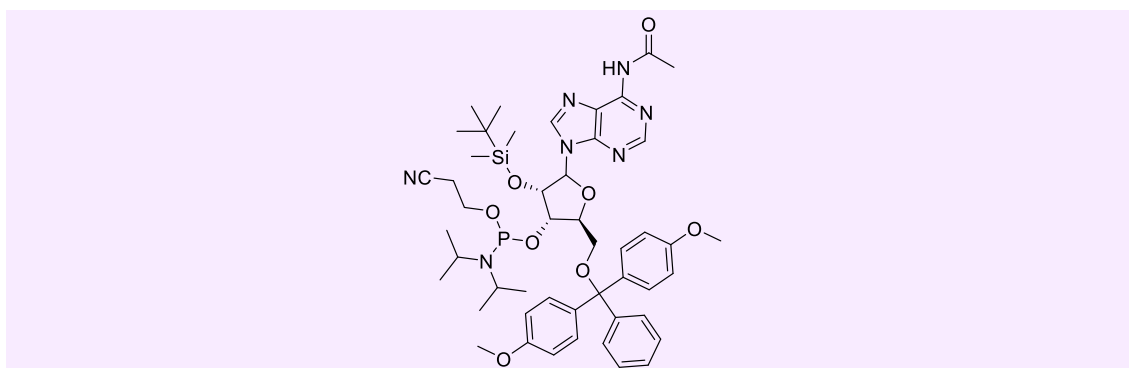
Figure 1.9. Structure of N2-Acetyl-2'-O-tert-butyldimethylsilyl-5'-O-DMT-guanosine 3'-CE phosphoramidite.

### 1.1.3.1. Lifetime of an Aptamer

Aptamers are better if their lifetime in biological environments is longer as this helps with therapeutic applications. One of the most effective ways of controlling the lifetime of DNA aptamers *in vivo* is to bind the aptamer with polyethylene glycol (PEG), which then prolongs their circulation in the bloodstream.[55] Other ways that have been successful to improve the life span are to coat the DNA aptamers in nanomaterials such as copolymers, liposomes and nanoparticles.[19]

### 1.1.3.2. Nuclease Degradation

As previously discussed, DNA and RNA aptamers are susceptible to break down by nucleases. Nucleases are found in the body and can cleave phosphoester bonds, which can destroy the structure of the nucleic acid. This can be resolved by chemically modifying the inter-nucleotide linkages and nucleotide sugars. Another way this can be done is by capping the ends of the chains or modifying nucleotides using the functional groups amino, o-methyl or fluoro.[56]

### 1.1.3.3. Renal Filtration

Most aptamers are commonly only 5-15 kDa molecular mass. Due to their small size, aptamers suffer high renal clearance, as the filtration cut-off of the glomerulous is 30-50 KDa, thereby decreasing bioavailability. To challenge this, aptamers can be joined to a high molecular mass polymer such as PEG, which can then stop aptamers being subject to metabolic processes and renal filtration.[57] This loss of the aptamers by renal filtration,  limits their therapeutic applications. PEG is an excellent tool for polymer-based drug delivery because it increases stability and solubility and also reduces toxicity and immunogenicity.[58] Macugen aptamer was modified with 40 kDa PEG, which increased its half-life in the blood circulation to 9 hours following intravenous injection from a time of 2 minutes before.[59]

## 1.1.4. Systematic Evolution of Ligands by Exponential enrichment

In 1990, A. D. Ellington and J. W. Szostak created a method that used combinatorial nucleic acid libraries to select RNA aptamers that bind selectively and tightly to a chosen target. There were studies on the interaction between the bacteriophage T4 DNA polymerase (gp43) protein and the ribosome-binding site of the mRNA, which encodes the enzyme. The gp43 binding sequence was selected from the RNA pool which was randomised at specific positions, a process called Systematic Evolution of Ligands by Exponential enrichment (SELEX). In 1992, a successful selection of single stranded DNA sequences from a chemically synthesised pool of random

sequence DNA molecules was shown. These DNA aptamers were able to identify and bind the ligands in the same way as the RNA aptamers.[60]

SELEX is also known as *in vitro* selection, a common procedure which consists of selections and amplification where large pools of nucleic acids (~$10^{14}$ molecules of up to 100 nucleotides in length[61]) are required to bind to selected targets under set conditions.[4] The general SELEX procedure is: binding, partition, elution, amplification and then conditioning.[60] In RNA SELEX, typically the library of single stranded RNA aptamers are prepared by the *in vitro* transcription of double-stranded DNA templates, generally using T7 RNA polymerase.[62] In DNA SELEX, the library of single stranded DNA molecules is mostly prepared by the separation of double stranded PCR products.[63] Firstly, a nucleic acid sequence is bound to the selected target and only the species capable of binding are then retained for the next step.[61] In both cases, the randomised DNA and RNA pool is incubated with the chosen target, the selection is designed to identify those molecules that have the greatest affinity and specificity for that target. The separation of the binding and non-binding aptamers is one of the most crucial aspects of the aptamer selection process. The most common way of separation is the use of affinity chromatography with immobilization of the target on the column material (agarose or sepharose commonly used). However, a large amount of target is needed for the column.[64] Another way is the use of magnetic beads for target immobilisation. This method is good because it requires only a very small amount of the target.[65] Other methods of separation include ultrafiltration using nitrocellulose filters with molecular weight cut offs,[66] capillary electrophoresis,[67] electrophoretic mobility shift assay[68] and flow cytometry.[69]

Next, the sequences retained are then amplified by PCR for DNA SELEX or by reverse transcription PCR for RNA SELEX. The resulting double-stranded DNA is converted into a new oligonucleotide pool by separating the significant single stranded DNA or by *in vitro* transcription and then purifying the produced RNA. This new and enhanced pool of selected

aptamers is used for a binding reaction with the target in the next round of SELEX (figure 1.10). The number of rounds is dependent on the design of the starting aptamer library, the reaction conditions, the features of the target, the concentration, the ratio of the aptamer and the target and the efficiency of the separating procedure. The number of rounds is usually between 5 and 15.[70]

Amplification is important because the library usually created is not very large because of the selection step and so amplifying them enriches the aptamer pool. The last SELEX round is stopped after amplification and the products are separated to give individual copies of the aptamers from the selected pool. Modifications can now be made to the aptamer sequences to enhance stability or for them to be experimentally used.[60]
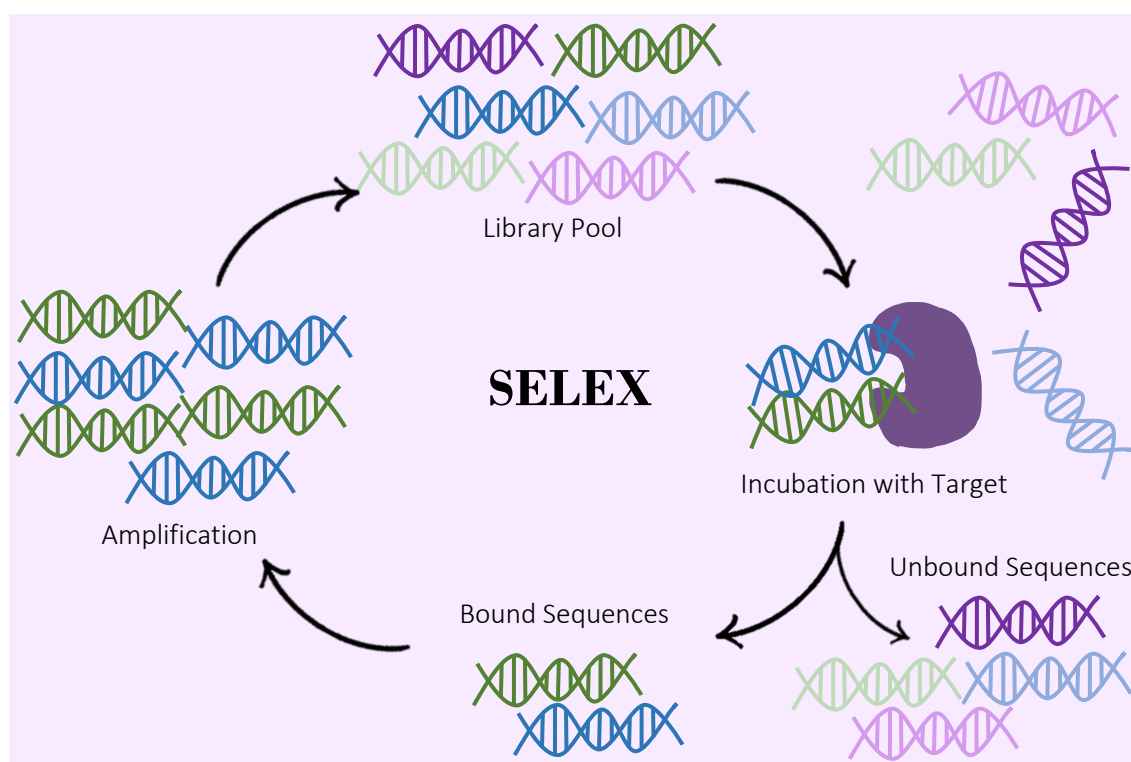


Figure 1.10. Cycle of Aptamer Selection.[71]

### 1.1.4.1. Importance of SELEX

SELEX has become a standard procedure used by research groups for generating aptamer pools. It is constantly being developed and new procedures added to make it more efficient. In 2018, a research group from Germany discovered aptamers for Protein A by applying next generation sequencing (NGS) to a selected aptamer pool which was obtained by SELEX using the FluMag-SELEX procedure followed by cloning and sanger sequencing.[72] FluMag-SELEX is a modified version of SELEX. It is modified by using fluorescent labels for DNA quantification and by the use of magnetic beads for target immobilization. These modifications are added to avoid radioactive labelling and the immobilization on magnetic beads enables easy handling, use of small amounts of target for the aptamer selection and rapid and efficient separation of bound and unbound molecules.[65] Sanger sequencing is the procedure of selective incorporation of chain-terminating dideoxynucleotides by DNA polymerase during *in vitro* DNA replication.[73] NGS is one of several high-throughput approaches to DNA sequencing using the concept of massively parallel processing. These technologies use miniaturized and parallelized platforms for sequencing 1 million to 43 billion short strands of 50-400 bases each per run.[74] NGS helps reduce the time needed for the SELEX aptamer selection procedure, focusing only on a few selection rounds. Through the use of SELEX, new sequence groups were found to have high specificity for Protein A.[75]

SELEX has been used along with a quartz crystal microbalance (QCM) to select aptamers that are specific to acrylamide. Acrylamide is a small molecule that potentially causes cancer, so finding an aptamer that can bind to it efficiently could be of great use in cancer research. This selection was done by starting with a single stranded DNA library ($10^{15}$ molecules of 45 nucleotides) performing 14 rounds of positive selection. The acrylamide aptamer candidates were enriched and two aptamer candidates were selected (A5 and C14) (figure 1.11).[76]

A5———TGGTCGTG—GTGAGGTGCGTG——TATGGGTGGTGG—ATGAG—TG——TGTGGC—————— 
C14————GCCGCCCCCAGTTGACCATGCCCACACCAT——GCCCTCA——C———CGCGCAC————

Figure 1.11. Aptamer Candidates A5 and C14.[76]

### 1.1.4.2. Slow Off-rate Modified Aptamers

Slow Off-rate Modified Aptamers (SOMAmers) are short single-stranded DNA-based protein affinity reagents selected *in vitro* from large libraries for their ability to bind to chosen targets. SOMAmers are unlike aptamers in that they have dU residues that are uniformly functionalized at the 5-position with different functional groups that can participate in interactions with target molecules and also form novel secondary and tertiary structures within the SOMAmer (figure 1.11).[77] Theoretically SOMAmers can be raised against multiple different biological substrates. They are called Slow Off-rate Modified Aptamers because they have very low dissociation rate constants.[78] They are identified using SELEX, the aptamer selection technology (section 1.1.4.1).[79] Scientists have given them protein-like properties by adding functional groups that copy the behaviour of amino acid side chains analogous to chemically modified nucleotides, which expands their chemical diversity and also gives them greater target variety as well as high affinity. SOMAmers are chosen because they have high affinity for their protein targets and also because they have low dissociation rates with these targets.[80]
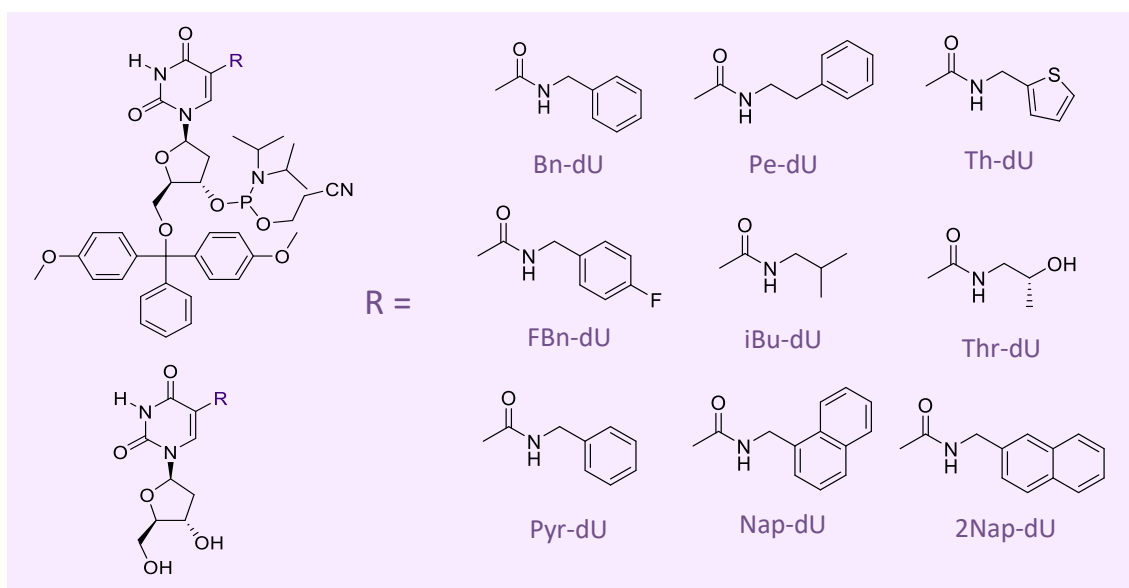
Figure 1.12. Structure of several modifications made at the 5-position of deoxyuridine available for SELEX and post-SELEX optimization.[78]

The hydrophobic modification made to SOMAmers regents on the 5-position of deoxyuridine nucleotides increases the resistance to degradation compared with other aptamers (figure 1.12). This, however, also depends on the nature of the modification at the 5-position and the sequence. Nuclease degradation can also be inhibited by such modifications as adding a phosphorothioate bond, which then substitutes a sulphur atom for a non-bridging oxygen in the phosphate backbone of an oligonucleotide. Most of the time this modification helps the internucleotide linkage too because this is more resistant to nuclease degradation. It has been recommended that three phosphorothioate bonds attached at 5' and 3' ends will inhibit degradation. The disadvantage to this modification is that it may also increase toxicity.[81] SOMAmers are claimed to have both high specificity and higher affinity for their protein targets and are less vulnerable to degradation by nucleases.[82] SOMAmer reagents are attractive for use in nearly every protein measurement application because of these great qualities along with their stability and easy production. The method to generate these SOMAmers is still very time

consuming, and researchers are currently working on new methods. This research project is the nest along from SOMAmers.

## 1.1.5. Aptamers Versus Antibodies

As aptamers have the ability to selectively interact with a particular target, they are often compared to antibodies. Monoclonal antibodies play a huge role in therapeutics and diagnostics. A large amount of research has been conducted into improving the performance of monoclonal antibodies for therapeutic uses. Generally antibodies are effective for most applications however aptamers are increasingly an attractive alternative.[83] In terms of selectivity, aptamers are easily selected against non-immunogenic molecules and molecules that are toxic to antibody producing cells. In terms of size aptamers are roughly ten times smaller than antibodies which allows them greater and easier access into cells and tissues.[84] Aptamers, unlike antibodies, can be selected under non-physiological conditions,[85] such as extremely high or low temperatures, or pH conditions. Aptamers can be readily chemically synthesised with minimal batch to batch variations as well as in large quantities, because of their smaller size compared with antibodies and the fact that many antibodies are still produced through biological routes in mammalian cells using recombinant DNA technology.[54] Aptamers provide complex folded tertiary structures with sufficient recognition surface areas to rival and even better antibodies. In addition, they are easily selected and can withstand organic solvent and show low immunogenicity.  These properties, shown in Table 1.2, demonstrate the huge potential that aptamers have in the biomedical field.[86]

Table 1.2. Aptamers vs. Antibodies.[87]

| | Aptamers | Antibodies |
|---|---|---|
| **Target Range** | Can be produced against almost anything | Limited to targets that produce an immune response |
| **Development Time** | Several weeks | Several months |
| **Synthesis and Manufacture** | Produced *in vitro* (no animals used)<br><br>Low batch-to-batch variability<br><br>Chemically synthesised | Produced in animals and then by recombinant methods<br><br>Poor batch-to-batch reproducibility<br><br>Expensive reactors required |
| **Handling and Storage** | Stable at ambient temperatures<br><br>No freezing or refrigeration needed | Refrigeration/ freezing<br><br>Easily denatured |
| **Stability** | Refrigeration/Freezing<br><br>Can withstand repeated round of denaturation without loss of function | Made from protein. Easily denatured losing structure, function under wide range of conditions. |
| **Shelf Life** | Long – several years it can be frozen<br><br>Can be reused in some applications | Limited to 6 months and single usage |

| *In vivo* complications | No intrinsic immune response | Can cause immune response if used as a drug |
|---|---|---|
| **Down Stream Applications/ Assay Development** | Functional under a broader range of conditions defined by the customer based upon the intended end use | Will only work under near physiological conditions |
| **Affinity and Specificity** | High affinity and greater potential for specificity due to the nature of the isolation process | High affinity and good specificity |

However, the development of clinical therapeutic aptamers is lagging far behind therapeutic antibodies,[88,89] which are still dominating the global medicine market place.[90] There are several factors causing this: the lack of medicinal chemistry support, the production costs, reluctance to break from conventional approaches, delay and suppression of clinical translation and distribution of therapeutic aptamers.

## 1.1.6. How are Aptamers Used?

### 1.1.6.1. Aptamers as therapeutics

Aptamers have the potential to treat illnesses and disease that currently have no cure. There are several ways in which aptamers can do this: act as a drug delivery agent, act as an agonist and finally act as an antagonist. Aptamers can work as drug delivery agents because the aptamer has such high affinity and specificity that it can target cells with high accuracy to deliver the desired drug. Aptamers can act as an agonist by activating cell-surface receptors. They can also be designed to can act as antagonists, by inhibiting protein-protein interactions or receptor-

ligand interactions.[7] For example, Centauri Therapeutics, is a UK-based biotechnology company focused on the discovery and development of novel aptamers to treat life threatening diseases. They are using aptamers to target specific pathogens such as antimicrobial resistant microbes and also cancer cells, to elicit an immune response in the patient, which will eliminate the threat.[91]

In recent research, aptamers have been isolated against therapeutically relevant targets to treat such diseases including human immunodeficiency virus (HIV), cardiovascular diseases, hepatitis C virus (HCV), neurodegenerative diseases, influenza and some cancer linked targets. The first aptamer approved for therapeutic use by the USA Food and Drug Administration (FDA), was in 2004 and called pegaptanib.  It is an RNA aptamer that consists of 28 nucleotides[92] and is used for macular degeneration (figure 1.13).[93] Pegaptanib (commercial name Macugen) is provided as a solution for injecting into the eye and is used to treat adults with the wet form of age-related macular degeneration to improve vision. It has a 2'fluorophyridine modification that does not interfere with amplification using reverse transcriptase.[54] Its target is vascular endothelial growth factor (VEGF).[94] This disease affects the middle part of the retina at the back of the eye, and is caused by abnormal growth of blood vessels.[95] However, monoclonal antibodies have been developed that are more effective clinically for this purpose.[96]

Aptamers can be used for therapeutic purposes in a similar way to antibodies. Antibody therapy is a form of immunotherapy that uses monoclonal antibodies to bind to certain cells or proteins. This is done with the aim to stimulate the patient's immune system to attack those cells.[97] No organisms are needed for the *in vitro* selection of oligonucleotides, a big advantage when manipulating molecules.[84] Aptamer therapeutics can be developed for intracellular, extracellular and cell-surface targets. Aptamers have advantages over protein therapeutics in terms of synthetic accessibility, modifications and size.[4] Targeting proteins in extracellular or cell-surface targets removes the need for the therapeutics to cross the cell membrane,[7] which

has been a challenge because of the protective lipid bilayer. Theoretically aptamers can be used

as therapeutics in any disease for which extracellular blockade of protein – protein interaction

is needed.[98]



Figure 1.13. Chemical Structure of Pegaptanib Aptamer.

There are other oligonucleotide-based therapies that are approved for therapeutic use or in

clinical trials. Currently, there are several antisense oligonucleotide-based drugs that have

entered phase III clinical trials such as Alicaforsen, ALN-TTR02, and QPI-1002.[99] Alicaforsen as

an example, was identified using SELEX technology, helps stop the production of ICAM-1 which

is a cell surface receptor involved in the process of inflammation (figure 1.14). In inflammatory

bowel disease, the immune system incorrectly activates inflammatory processes. In this

situation, ICAM-1 is highly expressed, and this correlates with inflammation in the bowel.

Alicaforsen binds to ICAM-1 mRNA, and then ICAM-1 expression is reduced. Clinical trials of

Alicaforsen have shown reduced levels of inflammation, a reduction in symptoms for the patient

and an improvement to the health of the tissue shown by endoscopy.[100]

Figure 1.14. Chemical Structure of Alicaforsen Aptamer.

### 1.1.6.2. Aptamers used for Drug Delivery

As well as serving as direct therapeutics, aptamers have also been developed to deliver drugs to the cytoplasm or the surface of human cells.[101] Aptamers can function as targeted drug delivery carriers that may increase the efficacy of a drug and also reduce the destructive side effects of other non-targeted approaches such as treatments for cancer patients like radiotherapy and chemotherapy. This aptamer-based drug delivery approach was explored with human prostate-specific membrane antigen (PSMA), which is a transmembrane protein that is linked with prostate cancer and is overexpressed on the surface of solid tumours.[102] This method of aptamer-based drug delivery has now been tried with nanoparticles. This works by having nanoparticles that contain therapeutic agents, which are then coated with aptamer recognition elements for aptamer facilitated internalisation and drug release.[103]

Another significant application of aptamer-based drug delivery systems involves developing strategies to treat infection with HIV. Viral entry of HIV into immune cells occurs when envelope glycoprotein gp120, a glycoprotein located on the exposed surface of the viral particle, encounters the CD4 receptor and CC-chemokine receptor 5 (CCR5) on the outside of the host cell.[104] Work has been done to develop a gp120 targeted aptamer that works by inhibiting the interaction between the gp120 and the CD4 receptor. The gp120 targeted aptamer can also act as a cell-type specific delivery agent for siRNA (Figure 1.15).[105] Research has shown that the

gp120 aptamer has been able to deliver anti-HIV siRNA into HIV infected cells and inhibit HIV activity *in vitro*. Further studies in a humanised mouse model demonstrated that administration of the gp120 aptamer-drug suppressed HIV replication *in vivo*.[106] This shows that aptamer–drug conjugates have great potential for treatment of HIV patients.



**HIV gp120 Aptamer-tat/rev siRNA Chimera**

Linker
Sense Strand
5'
3'
3'
5'
dTdT
Antisense Strand
RNA Aptamer HIV gp 120
Dicer substrace siRNA against HIV Tat/Rev
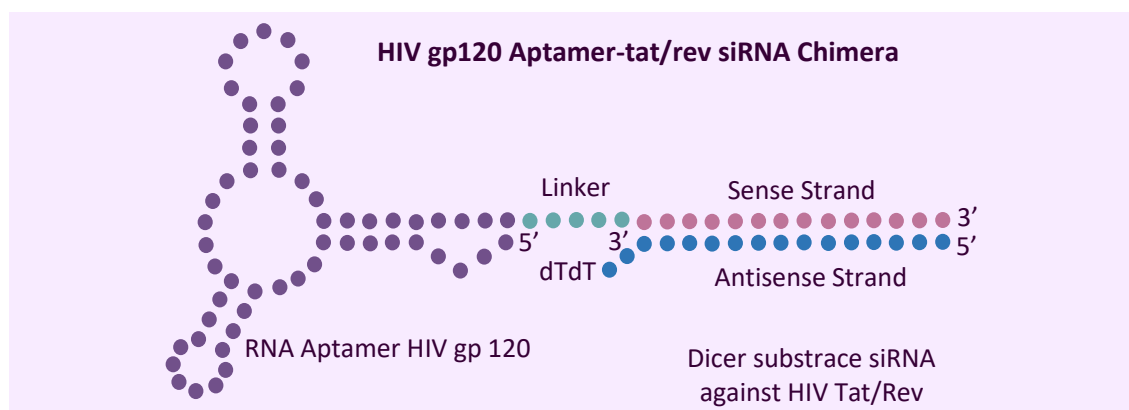
Figure 1.15. HIV gp120 Aptamer-tat/rev siRNA Chimera.[105]

### 1.1.6.3. Aptamers in Diagnostics

Aptamers could be selected for developing diagnostic tools for a range of diseases. This works as long as the appropriate targets such as pathogens or biomarkers are available. Signals can be generated by a conformational change in the aptamer, induced by the presence of a ligand which is ideal for the creation of sensors used in diagnostic applications. The ease of conjugation and labelling of aptamers allows them to be combined with other advanced technologies such as flow cytometry,[107] microfluidics,[108] cell separation,[109] endogenous nucleic acid analysis[53] and nanoparticle based sensing.[110] This makes aptamers a potential tool for diagnostics.[86] The key to effective treatment is an accurate and early diagnosis. Effective early diagnosis is only possible if the method is sensitive and involves small detecting agents, which is why aptamers are good candidates. Aptamers have been used in the detection of parasital, bacterial and viral-mediated diseases, along with cancer. Some of the existing diagnostic methods are in need of expensive equipment and skilled people, which when leads to high cost.[86]

Research has identified DNA aptamers that have been selected against cancer cells, which then can be utilised to diagnose and image cancer tissue.[111] This has been done with cells from liver, colon, pancreatic, prostate, and breast cancer, amongst others. This research has been helped by the use of the SELEX method.[112] Biomarker molecules that are specifically expressed or overexpressed on cancer cells were used for the selection of diagnostic DNA aptamers. Example biomarkers are vascular endothelial growth factor (VEGF),[113] prostate-specific membrane antigen (PSMA)[114] and epidermal growth factor receptor variant III (EGFRvIII).[115]

A DNA aptamer can be developed into a biosensor by chemical modifications with luminophores or linkage to nanoparticles to bind to a disease-related biomarker or a pathogen.[116] For example, research was conducted to develop a DNA aptamer, selected by SELEX and then PCR amplified to bind to and detect non-pathogenic Sterne strain Bacillus anthracis spores. This utilised an aptamer-magnetic beads-electrochemiluminesence sandwich assay scheme. At least three distinctly different populations of single-stranded DNA aptamers were identified. In the low DNA to spore ratio population, the aptamers were liberated from spore pellets by heating at 96˚ for 5 minutes after each round of SELEX. High affinity spore surface bound aptamers were detected by their 5'-biotinylated tails using a labelled avidin.[117]

## 1.1.7. Aptamer Limitations

The use of aptamers for applications is fraught with problems and limitations, which researchers are still working on to find solutions for. The main limitations are described below.

### 1.1.7.1. Aptamer Degradation

The rapid degradation of aptamers by nucleases in biological media, such as blood, is a large problem that puts limitations on their practical applications. Oligonucleotides can decay in blood within a few minutes to several tens of minutes, depending on the concentration and conformational structure of the molecule. This short time scale is undesirable in most

therapeutic applications, and because of this, several methods for protecting aptamers against degradation by nucleases have been developed. As discussed earlier (section 1.1.3), modifications such as 2'-amino pyrimidine,[118,119] 2'-fluoropyrimidine,[120] 2'-O-methyl purine and 2'-O-methyl pyrimidine[121] are used for this purpose.

### 1.1.7.2. Aptamer excretion from the bloodstream by renal filtration

Aptamer excretion from the bloodstream by renal filtration can complicate their therapeutic use. The kidneys can remove substances with a molecular weight below 30-50 kDa, and many aptamers are within this range. Conjugating aptamers with PEG with the molecular weight of 20-40 kDa is one of the most common solutions to this problem.[122] This method is currently being used to increase the circulation time in the blood stream of oligonucleotides, peptides and proteins.[123] The excretion of the PEG-conjugated aptamers from the bloodstream is much slower, taking up to several days, and they do not lose their specificity. These PEG-conjugated aptamers are more effectively delivered to tissues and organs.[124,125] Aptamers can also be conjugated with cholesterol molecules which prolongs their circulation in the bloodstream.[126]

### 1.1.7.3. Control of the duration of action

The pharmacokinetic parameters of a drug are essential to its therapeutic applications. The period of action depends on factors such as: degradation, their involvement in metabolic processes, renal excretion and many more. Many of these factors limit the drugs therapeutic applications. The use of aptamers as drugs or drug delivery vehicles can often solve the problems associated with controlling the period of action. One possibility is to synthesise a complementary oligonucleotide creating an antidote. The hybridisation of the therapeutic aptamer with the antidote then causes the aptamer conformation to change and it will lose the ability to bind to the target molecule. This approach has seen success in animal models, where an aptamer was delivered into the bloodstream and showed therapeutic effects. The antidote

was then injected into the bloodstream and the aptamer action stopped.[127,128] The high efficiency of aptamer hybridization with its antidote in the bloodstream offers a unique method to be able to control the duration of therapeutic action. This makes the application of aptamers preferable over antibodies as it is currently very difficult to control the duration of action of antibodies.

### 1.1.7.4. Interaction of aptamers with intracellular targets

Most aptamers have been selected using targets found on the cell surface or in the bloodstream. This theoretically makes their applications easy, as they just need to be delivered to the bloodstream to trigger therapeutic effects. However, recently there have been some advances in the intracellular delivery of aptamers. Special expression systems can generate aptamers inside cells and ensure they are accumulating in the cytoplasm or nucleus.

This can be done by transfection of cells with a recombinant vector expressing the aptamer sequence under a U6 promoter and has allowed specific inactivation of nuclear target proteins,[129,130] while aptamer expression under tRNA promoter control ensures predominantly cytoplasmic localisation of aptamers.[131] Cell-type-specific aptamer synthesis can be done by directional viral expression systems that deliver vectors to certain cells.[132,133] The concentration of the aptamers which are called intramers can be increased by using strong promoters and limiting the rate of aptamer degradation by nucleases through protection of the 3' and 5' ends with additional structures such as hairpins.[134] Aptamers can also be delivered to intracellular targets by the transfer of aptamers from the bloodstream to cells through receptor-dependent endocytosis.[135,136] Endocytosis of an aptamer binding to PSMA has been shown to provide specific and effective delivery of conjugated drugs to cancer cells expressing this antigen on their surface.[137,138]

### 1.1.7.5. Aptamer cross-reactivity

Aptamers have high specificity for their chosen target. However, they can also bind to other similar molecules. Four aptamers that were generated against DNA polymerase β have been shown to also bind to and inhibit DNA polymerase κ which belongs to another DNA polymerase family.[139] This aptamer cross-reactivity can be an obstacle to their therapeutic application because potential side effects can occur when the aptamers interact with other proteins. This problem can sometimes be avoided by adding a negative selection step in SELEX methods. This was proved by using SELEX to generate an aptamer that binds with high specificity for DNA polymerase ι and does not bind to DNA polymerases κ or to β.[140]

### 1.1.7.6. Automation of aptamer generation

The protocol for the generation of aptamers is relativity simple, although in real time it is a time and labour consuming process. The aptamer selection by SELEX does not always have the best affinity and specificity due to a sub-optimal SELEX procedure. New methods have been designed to reduce these limitations such as automated SELEX[141,142] which can be done in a few days and capillary electrophoresis SELEX (CE-SELEX). CE-SELEX includes a modified stage of selection. Non-equilibrium capillary electrophoresis of equilibrium mixtures is used for aptamer fractioning. This procedure takes 1-2 days and allows for more stringency for aptamers binding to a chosen target.[143,144]

## 1.1.8. Aptamer Modifications

### 1.1.8.1. Modifications made to Aptamers

Modifications have been made to aptamers to enhance their therapeutic suitability and to increase their half-lives. Some of the main problems with aptamers are that they are susceptible to nuclease degradation and renal filtration; this is caused by the small size of the aptamers and

if not modified can create short half-lives. For example, the half-life of an RNA aptamer in blood can be only 2 minutes long, and this can cause problems when using aptamers as therapeutics.[145] However, there are advantages to aptamers having short half-lives, as it means when aptamers are used as drugs in the body they do not need to have additional modifications to accelerate excretion from the body and they will not cause negative side effects.[57]

One of the most valuable modifications that can be made to aptamers for analysis is adding different fluorophores. These fluorophores can be added to either the 5' or 3' end of the aptamer allowing for the aptamer interactions to be monitored when they change. The target can then also be visualised if the aptamers have fluorophores attached, this allows them to be used more effectively than antibodies in methods such as western blotting. Because there are a large variety of fluorophores, many different targets can be looked at all at once on a single sample. Confocal microscopy can be used to visualise aptamers in tissues or be used to follow the binding of aptamers and targets. Aptamers that are labelled can also be used in fluorescence-activated cell sorting (FACS) analysis, which can identify particular cell types in a mixed sample.[146]

Light-up aptamers have been designed that allow the tagging of RNAs with a fluorescent probe: this provides a simple way of live-cell imaging of RNAs by fluorescence without using protein (figure 1.16).[147] These light-up aptamers being composed of RNA have great advantage over GFP, because RNA is highly modular[148] and its sequence can be quickly enhanced by cycles of transcription, selection and reverse transcription.[80,149,150] Light-up aptamers provide a direct measure of gene transcription at the RNA level and can  be used to give a more accurate observation of RNA localisation.[147] Light-up aptamers can now be used for RNA trafficking *in vitro*,[151] protein sensing RNA,[152] metabolite sensing via a transducer linked to metabolite-sensing aptamers,[153] ribonucleoprotein (RNP) purification[154] and ribozyme cleavage monitoring.[155]
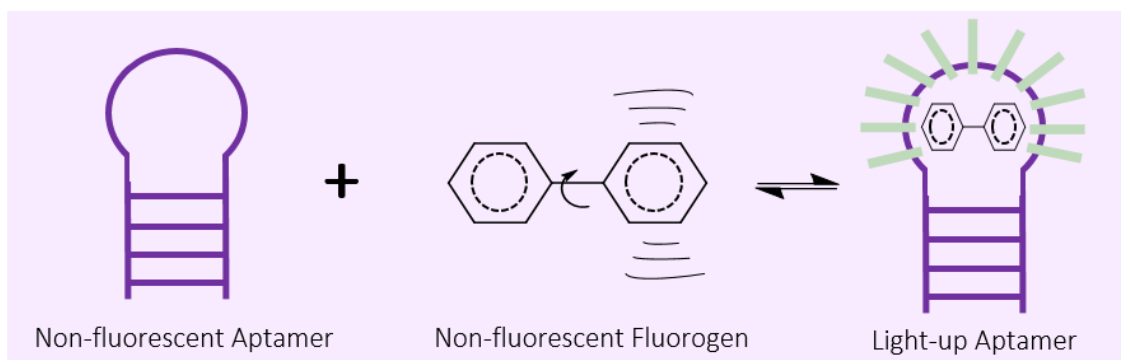
### 1.1.8.2. Aptamer Conjugation

If an aptamer is joined to another molecule, this can change the properties of the aptamer itself. Conjugation can be used to bridge the gaps between the target and the aptamer, and can help form additional bonds. Aptamers can be used as drug delivery systems when joined to drugs. This is a good method because aptamers have such high specificity and affinity that they bind only to their desired target and do not touch healthy cells and tissue.[56]

On a fluorescently labelled aptamer, a fluorophore is conjugated at one end and a quencher at the other end of the aptamer. If the aptamer is without its target it will fold into a hairpin confirmation which positions the fluorophore close to the quencher. When a ligand binds the base pair, interactions are interrupted and the aptamer goes back to the open conformation. The fluorescent signal is used for detection and quantitative measurement of the target concentration (figure 1.16).[156]

### 1.1.8.3. Evolution of sequence-defined highly functionalized nucleic acid polymers

David Liu and colleagues have designed a new genetic code to offer a high degree of codon and side-chain diversity to evolving polymers.[157] This has been done by a ligase-mediated DNA-templated polymerization system and *in vitro* selection system to create Highly Functionalised Nucleic Acid Polymers (HFNAPs) made from 32 building blocks of eight chemically diverse side chains on a DNA backbone (figure 1.17). It does so without the constraints that are added by

using polymerases or ribosomes. Through iterative cycles of polymer synthesis, selection and reverse translation, they have discovered HFNAPs that bind proprotein convertase subtilisin/kexin type 9 (PCSK9) and interleukin-6, which are two protein targets implicated in human inflammatory diseases.[158] This evolved polymer potently inhibited the interaction between PCSK9 and the low-density lipoprotein receptor. Structure–activity relationship studies revealed that specific side chains at defined positions in the polymers are required for binding to their respective targets. The DNA-templated, ligase-based translation system supports selection of polymers with diverse side chains, including side chains that mimic the repertoire of amino acid side chains found in proteins. PCSK9-binding and IL-6-binding polymers both exhibit position dependent and side-chain dependent structure activity relationships that mimic those of proteins. The PCSK9-binding polymers are strongly dependent on the presence of multiple side chains with different physical properties, consistent with the importance of chemical diversity to the functional potential of these polymers. The limitation to this method is that the variation is only in the side chains on particular nucleotides.[157]
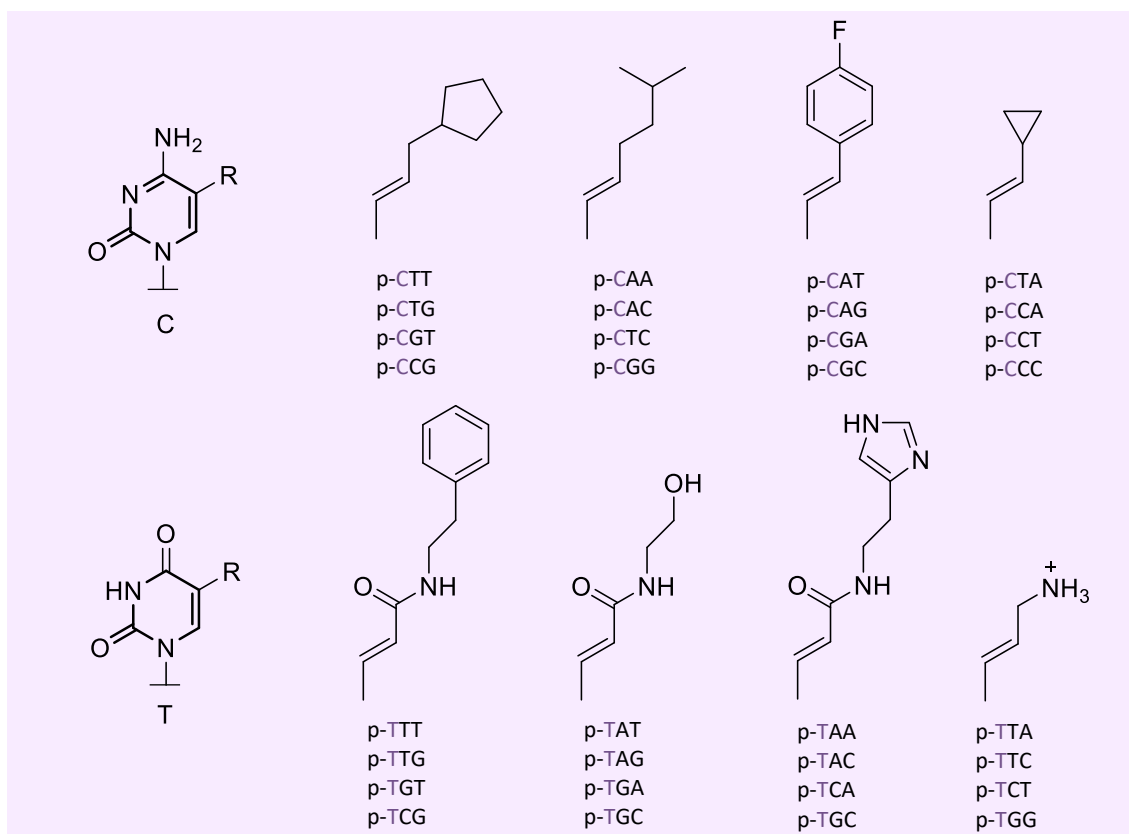
Figure 1.17. Structures of 5'-phosphorylated trinucleotide building blocks for HFNAP library synthesis.[157]

Gawande and colleagues have performed selections for PCSK9 aptamers from modified DNA libraries. These libraries had one or both pyrimidines altered to side-chain-functionalised variants.[159] High-affinity aptamers were enriched from doubly modified libraries where hydrophobic or phenolic side chains appeared on average on 50% of the nucleobases. Liu's findings are consistent with this; the highest affinity binder from their HFNAP library, roughly containing an equal mix of hydrophobic and hydrophilic side chains installed at 33% total frequency, has $K_d = 3$ nM for PCSK9. The flexibility of the side chains also allows for the possibility of creating evolution experiments with libraries with different side chains to discover the relationship between the structure of building blocks in a genetic code and the biological potential of the polymers. Other modifications may be suitable for other applications.

The evolution of these sequence-defined synthetic polymers made of building blocks beyond those compatible with polymerase enzymes or the ribosome, has the potential to generate new

classes of receptors, catalysts and materials. Liu's findings expand the chemical space of evolvable polymers to include densely functionalized nucleic acids with diverse, researcher-defined chemical repertoires. It enables researchers to generate and select HFNAPs with side chains tailored towards specific applications.[157]

# 1.2. Combinatorial Chemistry

## 1.2.1. Combinatorial Libraries

The synthesis of combinatorial compound libraries is being used more and more to lead to the identification of novel compounds that have targets with activities of high interest. These chemical libraries are intentionally made with a selection of different molecules prepared biosynthetically or synthetically and are be screened for a specific biological activity. There are a number of different library types: libraries of soluble molecules, peptide libraries on biological display vectors, libraries of compounds bound to resin beads, and other solid supports.[160]

Since the 1990's, combinatorial chemistry has become very popular in both academia and industry, because it can assist in drug discovery. It has been regarded as one of the most important new advances in medicinal chemistry.[161] Every combinatorial library method has three steps. Firstly the preparation of the library, secondly the screening of the library and finally the determination of the chemical structures of active compounds. These libraries involve synthetic chemistry and can include oligomers, oligonucleotides,[149] peptides,[162] synthetic amino acids and small molecules.

The idea of combinatorial chemistry was first developed in the mid 1980's with Houghten's tea bag technology to synthesise thousands of peptides on a solid support in parallel followed by Geysen *et al.* with multi-pin technology.[163] In 1991 Lam *et al.* introduced a combinatorial chemistry approach called the one-bead-one-compound (OBOC) method of combinatorial

peptide libraries[164] and Houghten *et al.* demonstrated the solution phase mixtures of combinatorial peptide libraries.[162] Bunin and Ellman reported the first combinatorial library of small molecules in 1992.[165] In a similar fashion, the phage-display peptide library method was designed in 1985 by Smith.[166] This method displays one-page one-peptide similar to OBOC library methods. Each M13 phage displays 5 copies of one unique peptide. In the 2000's, Frankel *et al.*,[167] Josephson *et al.*[168] and Murakami *et al.*[169] all described creating mRNA display macrocyclic peptide libraries using building blocks of unnatural and D-amino acids. In 2009, Heinis *et al.* presented the method of post-translational chemical modifications of phage-displayed peptide libraries.[170] These approaches enable peptide libraries to have greater chemical diversity and more resistance to proteolysis which makes them potentially better as drug candidates.

Recent advances have introduced DNA-encoded libraries (DEL) which allow researchers to create and decode very diverse small molecule, peptide or macrocyclic libraries.[171] This technology is a novel ligand identification strategy that allows the synthesis and screening of highly chemical diverse libraries more efficiently than the original method.[172] This was conceptually designed by Brenner and Lerner in 1992.[173] This method connects molecular biology and organic chemistry disciplines by using synthetic chemistry cycles to introduce diverse small molecules encoded by unique DNA tags.[174] There are then several cycles of selection, usually involving an immobilised target protein and a pool of libraries. This results in a mixture of compounds which are top binders of the target protein.[175] There is then amplification of the DNA region by PCR methods and subsequent next generation sequencing (NGS) allows for the identification of the binding molecules structure.[176] The progress over the last two decades has made DEL a powerful tool for pharmaceutical companies to identify new ligands for biological targets.[177, 178]

## 1.2.2. One-Bead One-Compound Libraries

The OBOC combinatorial library method is very efficient; millions of compounds can be screened thoroughly, simultaneously.[179] The concept is based on the idea that a combinatorial bead library can be prepared by the 'split and mix method' which results in a single bead carrying many copies of a unique compound.[180] This method has been successful in identifying ligands and a great number of biological targets. The OBOC combinatorial library approach has some key unique features. (1) The very quick synthesis of a large library of very high structural diversity is enabled by the 'split and mix' method. (2) All compounds can be tested and screened simultaneously but independently of each other against a target of interest because the compounds are spatially separable.[181] (3) This allows for rapid identification of the most active compounds whose chemical structures can be identified on the bead. (4) On-bead binding, functional assays and solution-phase assays can all be used in this method (figure 1.18). Once identified, the chemical structure of the compound on a positive bead may be determined directly or by an encoding strategy. The OBOC library method has not been used to its full potential due to technical obstacles such as screening millions of beads by conventional methods, for example by using a microscope to manually remove 'hits', which is unfeasible for a large library.
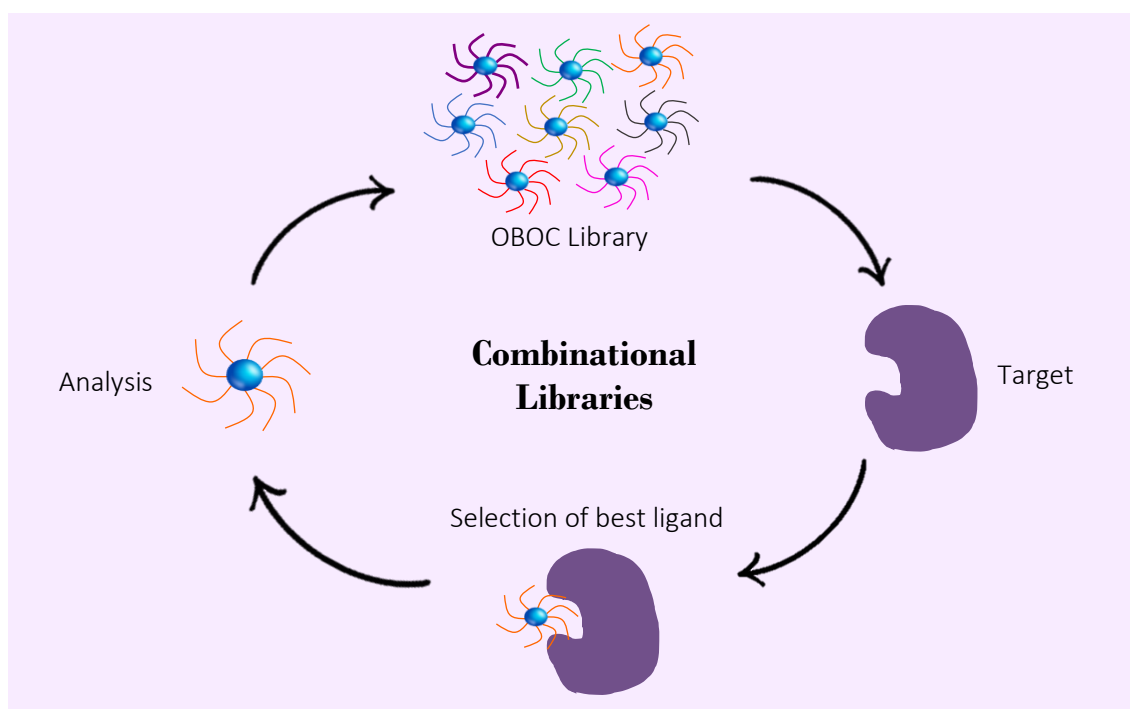
Figure 1.18. One-bead One-compound Combinatorial Libraries.

The OBOC library technique was first used successfully on over a million peptide beads used to isolate peptide ligands that interact with anti-α-endorphin monoclonal antibody using an enzyme-linked colorimetric detection technique.[182] Over the years, this OBOC technique has been successfully used on a number of different targets such as antibodies,[183] protein kinases,[184] bacteria,[185] G-protein coupled receptors,[186] whole cells[187] and small organic dyes. In recent research, Lam and colleagues used the OBOC method when looking into αvβ3 integrin ligands.[188] αvβ3 integrin ligands are of great interest because they have the potential to be used in cancer imaging and targeted therapy. αvβ3 integrin is expressed on the surface of numerous cancer and non-cancerous cells.[189] It is involved in various processes such as the formation of new blood vessels (angiogenesis), bone resorption, metastasis of human tumours and the death of cells (apoptosis), which is why it could play a key role in cancer therapy. This research team designed OBOC combinatorial libraries containing the arginine-glycine-aspartic acid (RGD) motif.[190] This library was then screened against K562 myeloid leukaemia cells that had been transfected with the human αvβ3 integrin gene. The team compared the binding strengths of

LXW7 (a cyclic peptide) ligand, the biotinylated form of LXW7 ligand and biotinylated RGD

cyclopentapeptide ligands to αvβ3. The cyclic peptide LXW7 was found to be the most suitable

ligand, binding to αvβ3 with an affinity of $IC_{50} = 0.68 \pm 0.08$ μmol/L. The biotinylated form of

LXW7 ligand produced a similar binding strength to the LXW7 against αvβ3 integrin and the

biotinylated RGD cyclopentapeptide ligands showed the weakest binding affinity. When further

research was undertaken, it was shown that the LXW7 ligand was also able to bind to the A375M

melanoma and to the U87MG glioblastoma cells. Both of these cell lines express high levels of

αvβ3 integrin. The research team also ran *in vivo* and *ex vivo* optical imaging studies in which

they delivered biotinylated LXW7/streptavidin-Cy5.5 complex in mice with A375M or U87MG

xenografts. This revealed that biotinylated LXW7 is preferably taken up into these tumours. This

was compared with biotinylated RGD cyclopentapeptide ligands and demonstrated that

biotinylated LXW7 has a higher tumour uptake but a lower liver intake. This study shows that

the OBOC combinatory library method is a powerful tool to discover and advance specific

ligands against cell surface receptors. The conclusion of the report shows that a peptide ligand

LXW7 with a RGD motif was discovered using one-bead-one-compound combinatorial

chemistry. LXW7 has potential to be an effective peptide ligand that can be used for drug

delivery.[191]

A recent research paper published in 2018 by a team led by Thomas Kodadek discusses the use

of OBOC libraries to screen serum samples to identify the ligands to antibodies in the mixture.[192]

This research was focused on the discovery of serum biomarkers for the diagnosis of disease

and for the prediction of drug efficacy. The immune system is an attractive source of these

biomarkers,[193] because disease states induce the production of antibodies against disease-

specific antigens. They have developed an approach called epitope surrogate technology to

discover these antibodies[194] using one-bead-one-compound combinatorial library,[188] with 'split

and mix' synthesis. The combinatorial library was then incubated with a pool of control serum
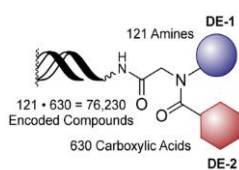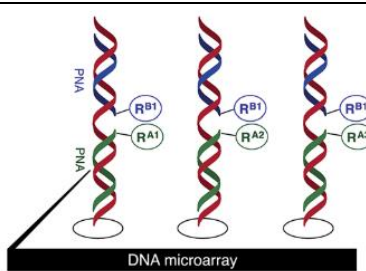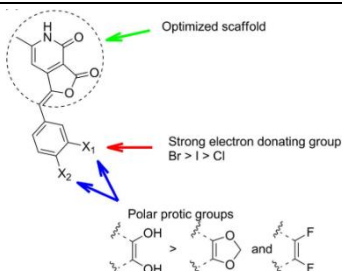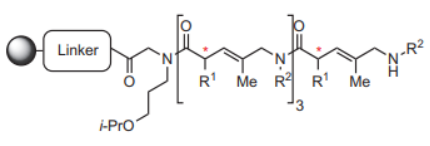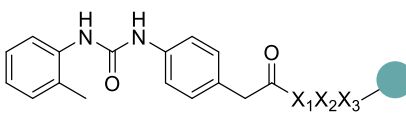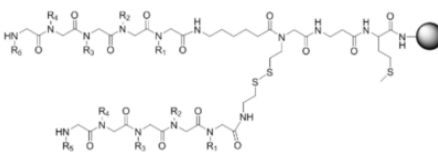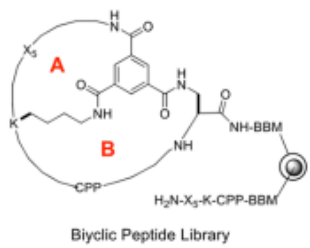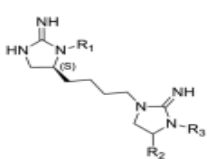
samples and then the 10 µM beads that show antibody ligands are stained with a fluorescently labelled secondary antibody. This is then repeated with the real serum samples. These beads can be passed through a flow cytometer and the fluorescent hit beads, which contain the ligand, are separated from non-ligand beads. All the beads that fluoresce contain the ligands that are potential antibody biomarkers, separated from the control population. Extra experiments can be run on these antibodies such as immunoprecipitation (a technique that precipitates a protein antigen out of solution using an antibody that specifically binds to that protein) to work out their native antigens in a suitable tissue extract.[195]

## 1.2.3. Combinatorial Chemistry in Drug Discovery

Over that last decade, combinatorial libraries have been successfully applied to many applications, including drug discovery. Table 1.3 below displays some of the different approaches and applications of combinatorial libraries.[171] Blakskjaer *et al.* describes a screening method named 'binder trap enrichment' which screens libraries homogenously.[196] Building blocks are spatially confined in the centre of the DNA junction which they have named Yoctoreactor, which facilitates both the chemical reaction between the building blocks and the library encoding. This screening can be performed in a single binding tube. This method has become an increasingly popular technology for the identification of small molecules targeting proteins.[171]

Table 1.3 Examples of recent application of combinatorial chemistry for drug discovery.[171]

| Library Type | Library Structure | Screening Type |
|---|---|---|
| DNA Encoded Chemical Library |  | Affinity Screening[197] |

| | | |
|---|---|---|
| Peptide Nucelic Acid-encoded small-molecule library |  Diverse array from combinatorial pairing of small molecule fragments $A^{1-125} \times B^{1-500} = 62\ 500$ combinations | Affinity Screening[198] |
| Spatially addressable solution-phase library |  | Plasmid Relaxation Assay[199] |
| OBOC Chiral Oligomers of Pentenoic Acids (COPA) Library |  | Protein-binding Assay[200] |
| OBOC Peptidomimetic Library |  | Cell-binding assay for a4b1 integrin ligand (LLP2A)[201] |
| OBOC Peptoid Library |  | In situ releasable assay against Cryptococcus neoformans[202] |
| OBOC Bi-cycle Peptide Library |  Biyclic Peptide Library | On-bead enzyme-linked assay against oncoprotein K-Ras G12V[203] |
| Positional Scanning Small-molecule Library |  | Antibacterial activity against the ESKAPE pathogens[204] |

# 1.3. Flow Cytometry

## 1.3.1. Flow Cytometry

Flow cytometry is a technology that rapidly analyses single cells and particles as they flow past several lasers, in a buffered salt-based solution. Each particle is analysed for visible light scatter and fluorescence characteristics. The visible light scatter is measured in two different directions; Forward Scatter (FSC) provides the relative size of the cell or particles, and Side Scatter (SSC) displays the internal complexity of the cell or particles. This light scatter is independent of the fluorescent intensity. Flow cytometry is a powerful tool that has interdisciplinary applications. These consist of cancer biology,[205] immunology,[206] virology,[207] molecular biology, infection disease monitoring[208] and cell drugging analysis. It is very effective for the study of the immune system and its response to infectious diseases and cancer. It allows for simultaneous characterisation of mixed cell populations that can be separated into single cell populations.[209]

## 1.3.2. Fluorescent Activated Cell Sorter

A fluorescent activated cell sorter (FACS) is a specific type of flow cytometry instrument which is used to analyse the chemical and physical characteristics of particles in a liquid as it passes through one or more lasers. Importantly, it can also purify and select out cells and particles for further analysis. Cells are labelled fluorescently and then lasers are used to excite the fluorescent label either on or within the cells, which then emit light at varying wavelengths. This process can sort a heterogeneous mixture of cells into different containers depending on the fluorescent characteristics and the specific light scattering of the cells. The cell sorter allows the user to select and gate off a particular population of cells or particles which is positive or negative for the desired parameters and then direct those cells or particles into the chosen

container.[209] This instrument is very useful because it is fast and provides quantitative and objective monitoring of fluorescent signals of cells.[210]
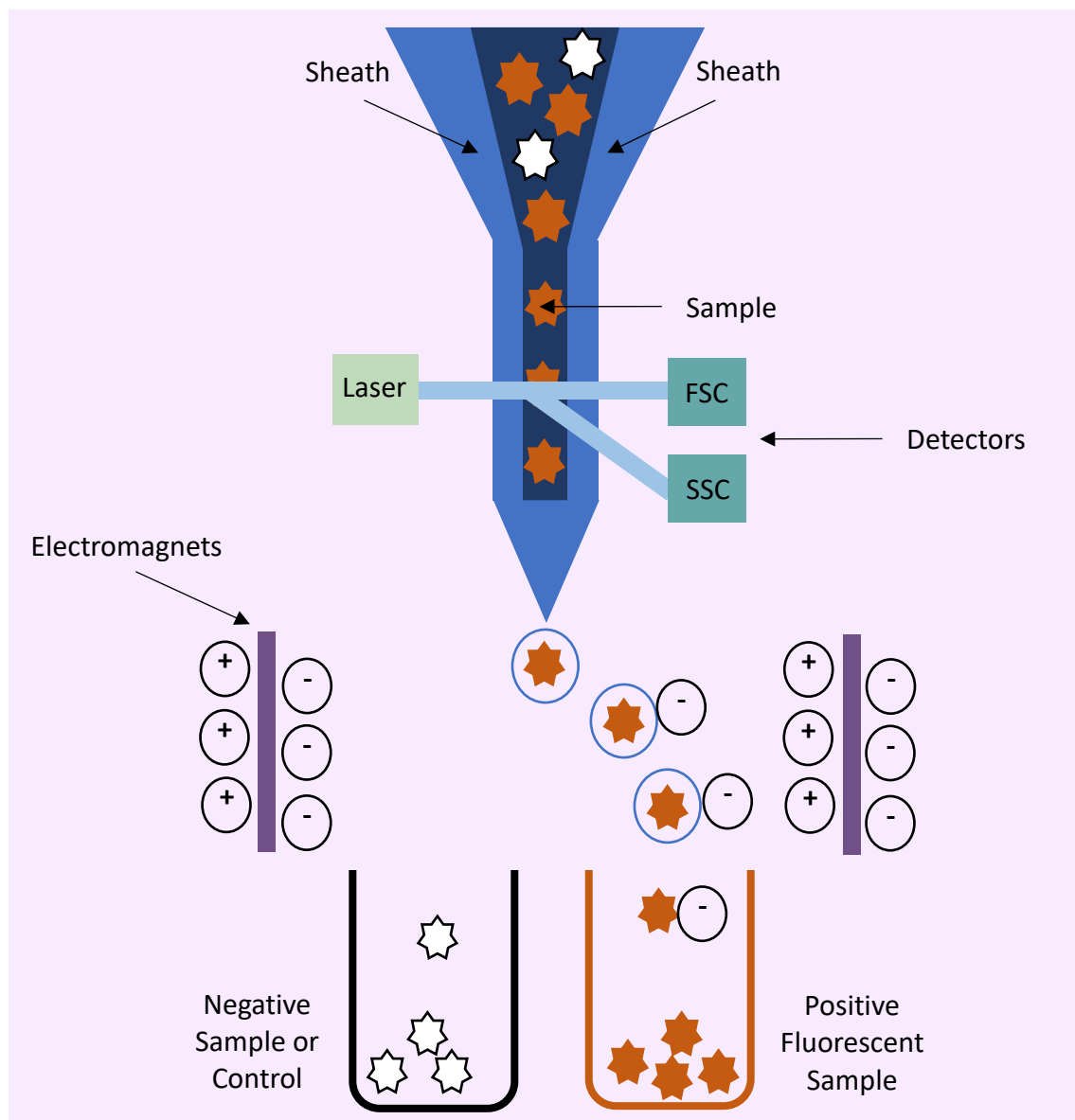


Figure 1.19. Flow Cytometry Schematic.

### 1.3.2.1. Hydrodynamic Focusing

A FACS instrument works by having the cell suspension in a narrow stream. The sample flow is focused such that the cells pass one by one through the laser beam. The stream is comprised of a constant flow sample stream and a constant flow sheath fluid (buffered salt-based solution) stream. Both streams are introduced into the microfluidic channel flow down either side of the

channel. The streams stay this way if they are coordinated on viscosity and hydrophobicity. The sheath fluid stream pushes on the sample stream to make it narrow; if the pressure is increased the sample stream widens. The sample stream is affected by sheath speed and conical tapering of the cuvette. The stream should be ~10 microns in diameter. The flow is arranged so that there is an effective separation between cells relative to the cell diameter.[211]
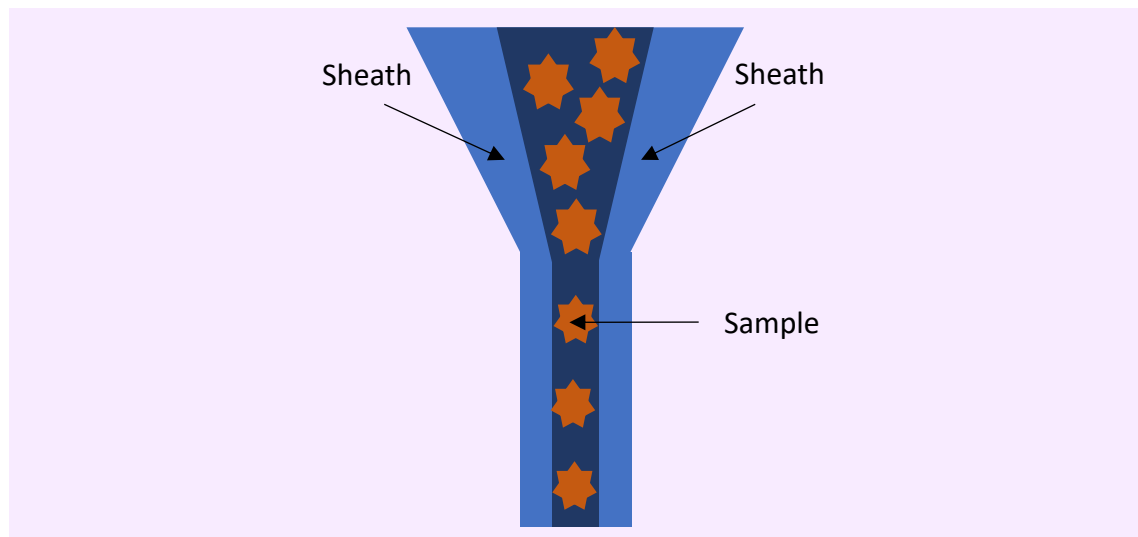


Figure 1.20. Hydrodynamic Focusing in the FACS.

Streams stability is essential to the precision of sorting cells. Many things can disrupt the stream such as drop delay (if incorrect will lead to the wrong event being charged at the break off point), temperature changes and blockages in the lines. Larger nozzle sizes can produce less stable streams and stream characteristics can change over the course of a long sort.

### 1.3.2.2. Detectors

Key parameters which this instrument measures include forward scatter, side scatter and fluorescence emission signals (figure 1.21). Forward scatter is the light that is refracted by a cell in the forward direction; it is typically 20˚ off from the laser beam's original axis. The FSC detector measures scatter along the path of the laser and uses that to approximate the size of the cells or particles. Usually, bigger cells will produce more forward scatter than smaller ones and so a stronger signal. Side scatter is the light that is refracted by cells and travels in a different

direction to the original path. It is measured at a 90˚ relative angle to the excitation line of the laser. This detector can provides information on the type of cell and its cellular complexity, e.g. T cells can be distinguished from B cells in splenocytes.[212] It is affected by shape and size of cells and sensitive to cytoplasm, membranes, nucleus. This allows FSC and SSC to be used to distinguish between different cell populations based on their type, liveness and size.[213]
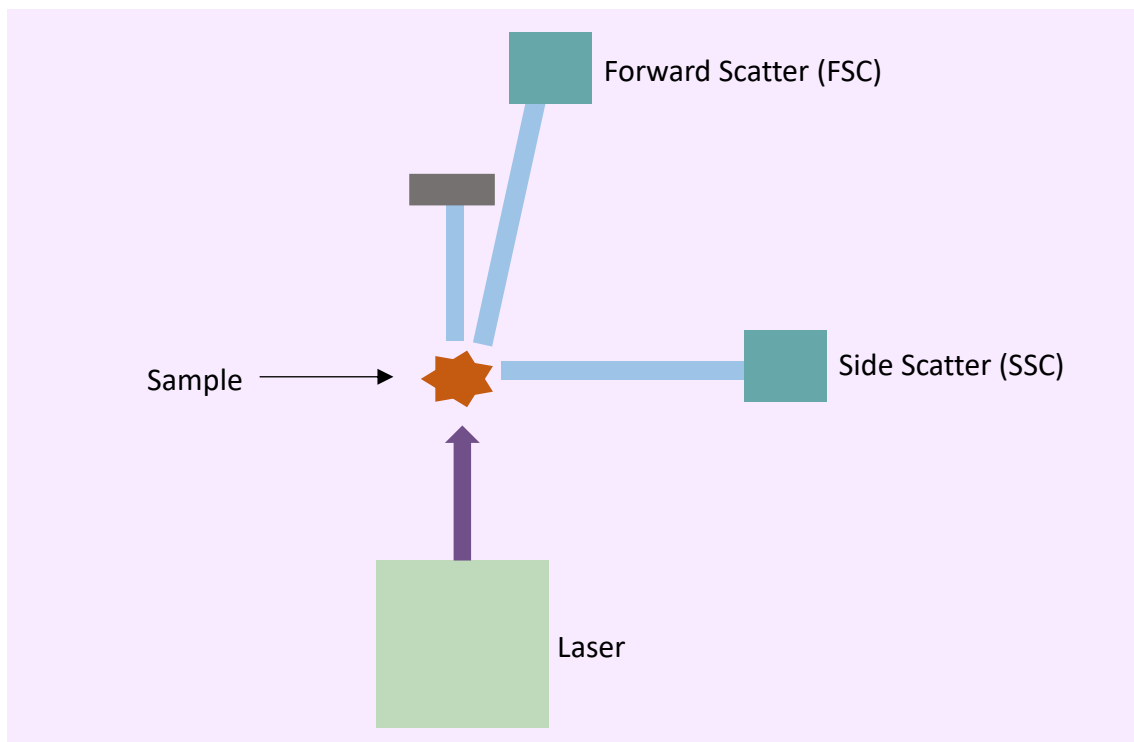


Figure 1.21. Detector on the FACS.

### 1.3.2.3. Lasers

The FACS can contain one or several lasers. A common three are those with excitation wavelengths of 488 nm, 405 nm and 561 nm also known as blue, purple and green lasers respectively. Each laser has filters at different emissions in the range of the excitation. Cell populations can be sorted based on their FSC and SSC, but cells can also be sorted by their fluorescent signal (figure 1.22). The fluorophore used for the detection emits light after excitation by a laser of compatible wavelength. Each different type of fluorescent dye or tag has its own characteristic excitation and emission level which is important for designing flow

65

cytometry experiments. Fluorescent samples can be prepared for flow cytometry analysis through several methods, such as transfection and expression of fluorescent proteins, for example Green Fluorescent Protein (GFP),[214] staining with fluorescent dyes such as propidium iodide or with fluorescently conjugated antibodies such as CD3 FITC, or chemically binding a fluorescent tag like rhodamine B to a particle.[209] There is a large range of fluorophores available: the Alexa Fluor family of fluorophores, allophycocyanin (APC), cyanine5.5 NHS ester (Cy5.5), fluorescein isothiocynate (FITC), peridinin-chlorophyll-protein (PerCP) etc. When the fluorescent cells pass through the interrogation point and interact with the laser, they create a pulse of photon emission over time. They will then be detected by the photomultiplier tubes and converted by the electronics system to a voltage pulse which is then registered as an event. The total height and area of the pulse is measured, and this directly correlates to the fluorescent intensity for that event.

When planning flow cytometry experiments there are key features to consider when picking a fluorescent tag: maximum excitation wavelength, maximum emission wavelength, extinction coefficient (the capacity for the fluorochrome to absorb light at a given wavelength), fluorescence quantum yield (the number of photons emitted per absorbed photon), fluorescent decay time (time interval after which the number of exciting fluorescent molecules is reduced), brightness, stokes shift (the difference between the maximum emission wavelength and the maximum excitation wavelength), which laser line to use, filter sets available and photostability.
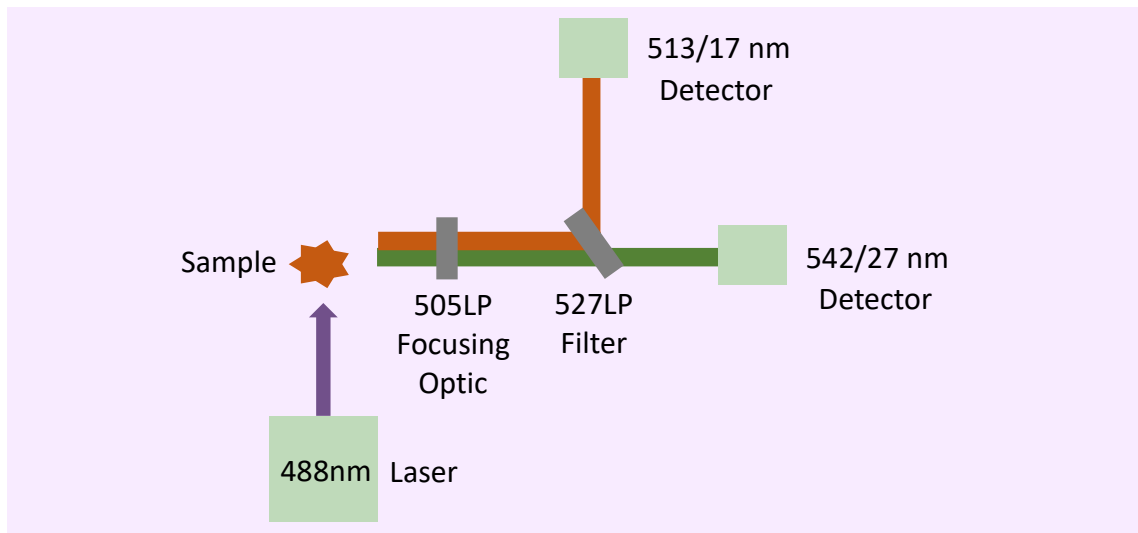
### 1.3.2.4. Vibrating Mechanism

The cell sorter separates cells by oscillating the sample stream of liquid at a high frequency to generate drops. At the place where the stream breaks into droplets an electrical charging ring is inserted, giving the droplets either a negative or positive charge. Charged droplets then go through an electrostatic deflection system which directs the droplets into a specific container based on their charge.[215] The containers can be tubes or plates (96-well plates are common).

### 1.3.2.5. Sample Preparation

The most essential requirement for all flow cytometry practices is to have the cells or particles in a single-cell suspension.[206] It is very important to obtain this to avoid blocking the flow system with clumps of cells. The concentration of the sample is important, if too concentrated the purity will be compromised, if too dilute the sort rate will be very low. Most samples are prepared in sheath fluid (bdbiosciences.com/sheath-fluid).

1.3.2.6. Data Analysis

The data for each event are plotted independently to represent the signal intensity of light detected in each detector for every event.[216] Data can be visually represented in several different ways, most commonly: contour diagrams, density plots, dot plots and histograms.[217]

1.3.2.6.1. Histograms

Histogram plots only measure one parameter. The y-axis is usually the event number that shows a level of fluorescence, and the x-axis is the relative fluorescent intensity detected for a single laser with a certain filter (figure 1.23). Many events that are detected at a specific intensity will be represented as a peak on the histogram. Flow analysis is often performed on a mixed population of cells and so produces multiple peaks on the histogram. This makes it important to have a negative control so that the positive data can be identified by the shift in intensity between the negative and positive.
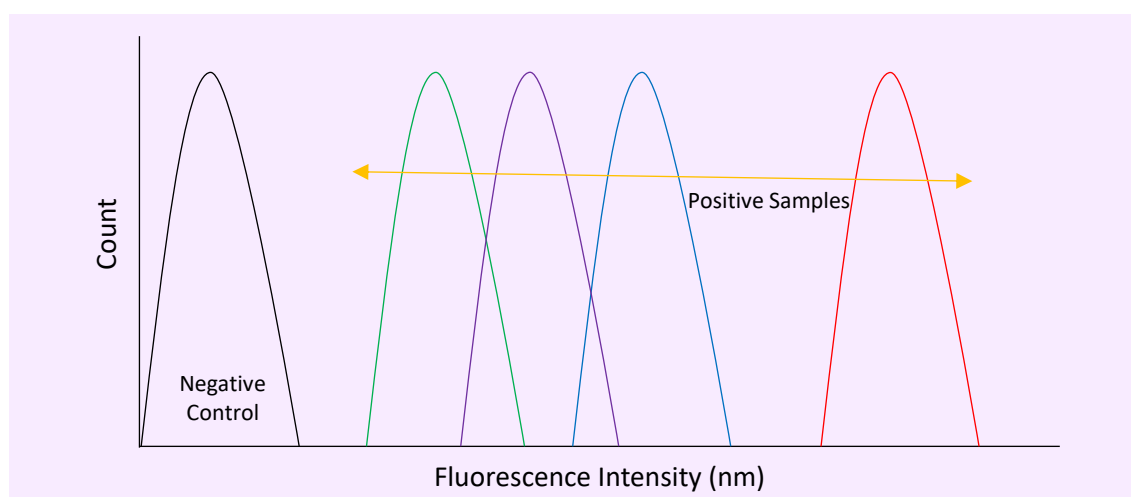


Figure 1.23. Histogram Graph showing positive samples against negative controls.

Sometimes the peaks on a histogram can appear as a doublet, this usually occurs when two cells or particle passes through the laser beam at the same time (figure 1.24). It can also sometimes be caused by having the voltage too high, amplifying the baseline. These doublets can

potentially have negative effects on the results, especially when sorting. The area and the width of the doublet peaks are larger than that of a single cell, this is because two cells will spend a longer time passing through a laser beam than one cell, but the height of the two peaks are very close. A way to overcome this is to run a doublet discrimination before running for samples or if your samples have a tendency to stick together than add in some extra steps to get a single suspension before analysis.[218]
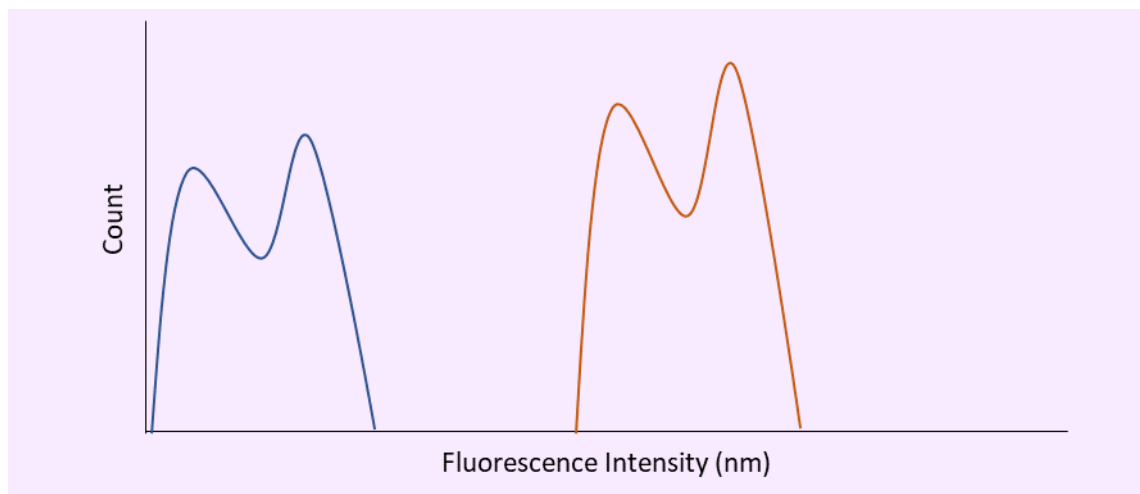


Figure 1.24. Histogram Graph showing doublet peaks caused by two cells or particles going through the laser at the same time.

### 1.3.2.6.2. Dot Plots and Density Plots

Dot plots and density plots often compare 2 parameters instantaneously where each event is represented by a single dot (figure 1.25). The dot plot shows the relationship between different variables against each other, for example FSC vs SSC or 513/17 nm vs 548/29 nm. The density plot was developed to show the relative number of events in a given region. Intensity measurements of different channels are represented along the axes showing that events with similar intensities will group together in the same region of the plot. Density plots are useful for observing rate of occurrence of sub-populations.
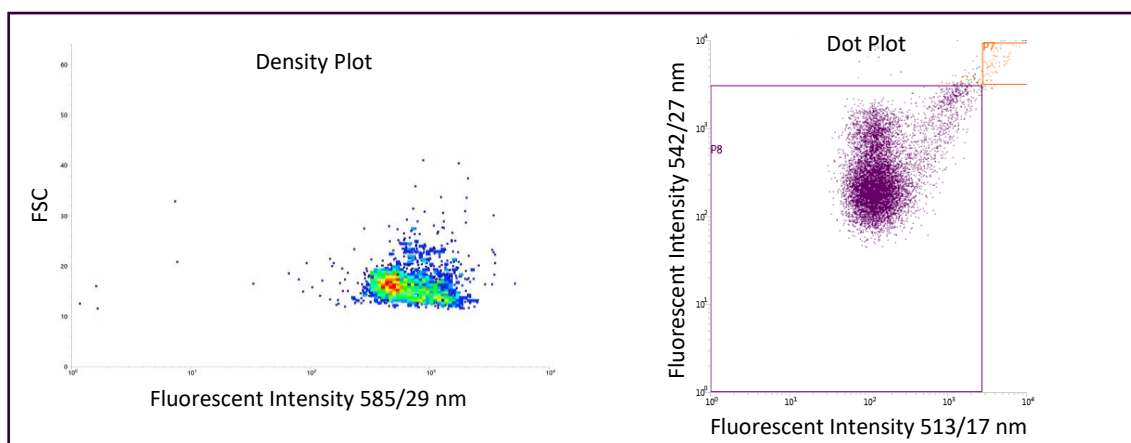
Figure 1.25. FACS Density and Dot Plots. Density is used to see how similar your population is and to spot outliers from the main populations. Dot plots are mostly used for sorting, to gate correctly the populations you want.

## 1.3.3. Flow Cytometry for Drug Discovery

In drug discovery, flow cytometry has been used at many stages.[219] Being a high content assay technology providing a large amount of information with multi-parametric analysis makes flow cytometry an important tool for drug discovery.[220] The development of faster and automated sampling processes to be used with flow cytometers has made flow cytometry more attractive to drug discovery and screening platforms.[221] Over the past decade, various flow cytometry screening applications have been undertaken, for example screens with yeast,[222] ligand-binding assays for screening novel G-protein-coupled receptor ligands,[223] and phenotypic screening, for example in a 1536-well plate of cytotoxic T lymphocyte lytic granule exocytosis.[224] As well as cell applications, bead based immunoassays have been performed to screen secreted cytokines and proteins in a multiplexed format.[225,226,227] Flow cytometry has many advantages over imaging based screening. For example, it can be used to analyse and physically separate individual cells from a mixed population without having to develop complex data analysis procedures, required with image-based screening. Flow cytometry is also optimal for screening suspension cells as well as adherent cells, after they have been detached from flasks or plates, however imaging

assays are only really suitable for adherent cell screening.[228] Below are examples of High-throughput screening using flow cytometry.

### 1.3.4.1. High throughput Flow Cytometry to identify inhibitors of the Rac1 GTPase

AstraZeneca has in place a large range of high throughput flow cytometry applications. This consists of High throughput screening (HTS), multiplex profiling and phenotypic to structure activity relationship (SAR) screening.[220] HTS is the principal hit-finding strategy at AstraZeneca. Essential factors in generating a successful screen are the quality, size and content of the compound library being screened. Safety-related concerns are the lead cause of drug discovery projects failing. Between 2005-2010, 82% preclinical drug discovery projects closed because of the unacceptable level of toxicity in compounds.[229] AstraZeneca worked to identify small molecules that act as inhibitors of the Ras-related C3 botulinum toxic substrate 1 (Rac1) GTPase to help prevent allergic airway hyper responsiveness (AHR). Asthma is a heterogeneous inflammatory disorder in which airways narrow and swell, this producing symptoms of difficulty breathing, coughing and wheezing.[230] It is a major public health problem as 5.4 million people in the UK are currently receiving treatment for asthma,[231] and asthma affects 262 million people worldwide.[232] AHR is the excessive contraction of the airway muscle cells, it is also seen in chronic obstructive pulmonary disease, which has limited treatments available.[233] A 348-well plate bead based flow cytometry assay for small GTPases has been published previously.[234] AstraZeneca worked with the biology company Intellicyt to reproduce this assay in a 1536-well plate format, as this is both cost- and time-effective. Beads coated in Glutathione were used to capture Glutathione S-transferase (GST)-tagged Rac1. Bodipy-labelled guanosine-5'-triphosphate (GTP) was used to detect the compounds that were competing with its binding to the GTPase which would then result in a decrease in fluorescence signal. Using iQue1Screener HD flow cytometer with laser excitation at 488 nm and emission at 530/30 nm, it was possible to detect the bead population, identify single beads and work out the fluorescence median on

the beads. The plates were then sealed and read (0.5 seconds sip time with 6 mL/well). The first screen was performed at 12.5 mM against a wide range of diverse compounds (500,000). The hits were tested in a ten-point concentration response (100–0.003 mM) against Rac1, along with the related GTPases Ras homolog gene family member A (RhoA) and cell division control protein 42 homolog (CDC42) for selectivity. Multiple compounds were selective for Rac1. A further panel of cellular assays is currently underway. Figure 1.26 shows dose-response curves of the three top hits, Rac1 (circles), Rho (squares), and CDC42 (triangles). These compounds were tested in a ten-point concentration response from (100-3 nM). The target coated-beads, bodipy-GTP and all 3 compounds were incubated together for 2 hours at room temperature. The final concentration of fluorescent GTP was 100 nM.[220]
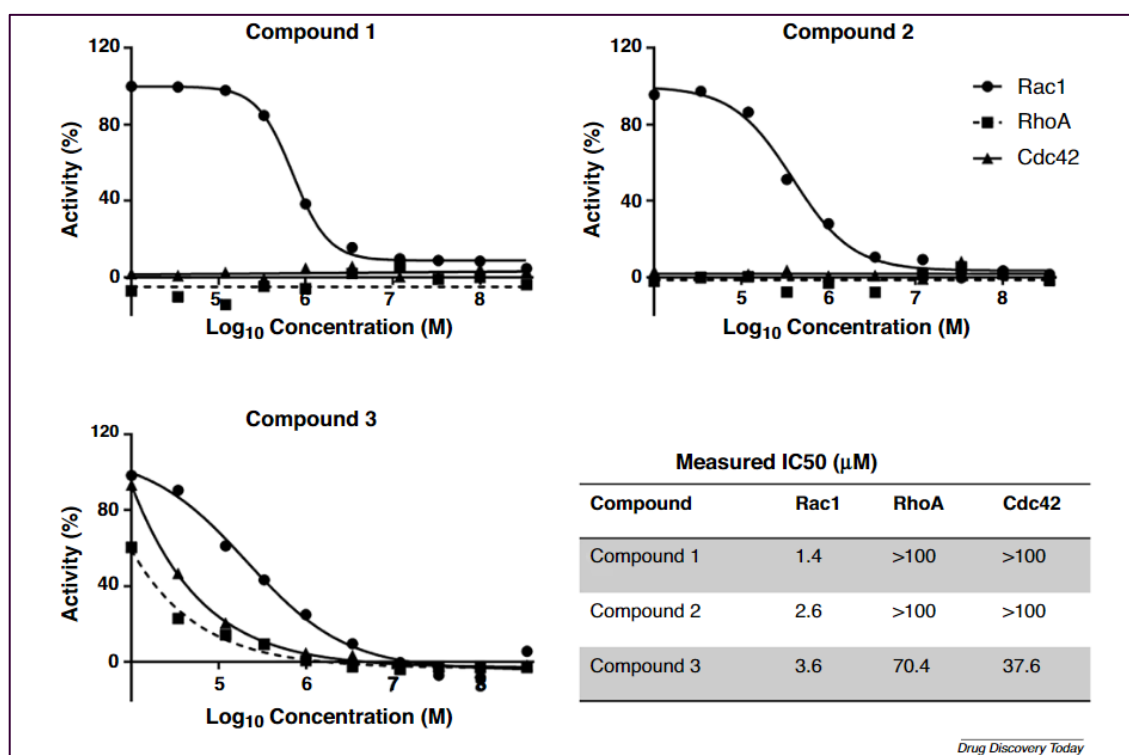


Figure 1.26. Dose-response curves of top three hit compounds.[220]

### 1.3.4.2. Directed Evolution of Aptamer Discovery Technologies

In 2022, H. Tom Soh *et al.* published a manuscript where flow cytometry was used to efficiently isolate high affinity aptamers in fewer selection rounds than the standard SELEX methods.[235]

This thereby minimised selection bias and reduced the excess material entering the final aptamer pool. They did this by having solution phase aptamers on particles and screening them by FACS to quantitatively isolate individual aptamer particles based on their affinity. They had a second method where they used a multiparametric particle display method. This used two-colour FACS to isolate aptamer particles with high affinity and specificity. This produced aptamers that demonstrated excellent target binding to protein glycoforms even in complex mediums such as serum.[235] This method is similar to the design of the project presented in this thesis, which was planned before this paper was published. The success of this method demonstrates that the use of FACS for identifying novel drug candidates can help drug discovery processes.

## 1.3.4. Flow Cytometry in Diagnostics

Flow cytometry can be used in clinical settings to characterise specific diseases.[236] Fluids such as bone marrow aspirate, cerebrospinal fluid and peripheral blood can all be analysed by flow cytometry. However, if the sample does not contain viable cells, then the analysis will not be successful. It is most commonly used for malignant and benign hematologic samples and can aid with diagnostics, treatment plans, monitoring progression of a disease or the success of treatment.[237, 238] Measuring levels of DNA content in a sample was one of the earliest uses of flow cytometry for diagnosis in 1983. If there was a 67% increase in DNA, the cells were defined as being malignant compared to non-malignant cells.[239] Malignant cells also show abnormalities in their chromosomes and this can also be quantified by flow cytometry.[240] The DNA abnormalities can then be retrospectively studied comparing their survival duration in patients.[241] Phenotyping is another common use for flow cytometry in oncology. There are many phenotypes used to differentiate between cancerous cells and healthy cells. Tumour cells often express distinguishable surface receptors.

Flow cytometry has also been used to examine the immune function of the carers of Alzheimer's disease patients.[242] It has been used to examine natural killer and T-cell function in patients with irritable bowel syndrome in correlation to psychological distress[243] and to examine the immunosuppression of women during pregnancy and the immune state of postpartum women over time.[244]

# 1.4. Cancer

## 1.4.1. What is Cancer?

In the UK, the predicted number of cancer caused deaths in 2021 has reached 176,400 with age-standardised rates being 89.2/100,000 for women and 114.6/100,000 for men.[245] In the European Union, the predicted number of cancer caused deaths in 2021 was 1,267,000. Corresponding to age standardised rates of 81.0/100,000 for women and 130.4/100,000 for men.[245] The American Cancer Society Estimates that there were 1,898,160 new cancer cases and 608,570 cancer deaths in the United States, in 2021.[246] There is some uncertainty in the predictions related to the effect of the coronavirus disease 2019 (COVID-19) which may have caused an increase in cancer mortality because of delay in diagnosis and interruption to treatment.[247,248] Cancer mortality worldwide continuously increased during the 20th century until 1991, after which it declined through to 2018 by 31%.[246] This was due to the reduction in smoking and improvements in early diagnosis and more successful treatments. This amounts to 3.2 million fewer cancer deaths than would have followed if the trend had persisted.[249]

Cancer is a collection of diseases in which normal cells become abnormal, exhibit uncontrolled growth and division of cells and acquire the ability to spread to other parts of the body, a process known as metastasis (figure 1.27).[250] The characteristics that distinguish cancer cells from noncancerous calls are the lack of differentiation, tendency towards tissue invasion,

metastasis and uncontrolled division.[251] These cells then start to become increasingly resistant to the controls that maintain normal tissue, and because of this they divide faster and become less dependent on signals from other cells.[252] Cancer cells can evade cell death which occurs as a normal and controlled part of an organism's development and growth, despite the fact that they are a prime target for apoptosis. In the later stages of cancer, cells break through normal tissue boundaries and metastasize to new sites in the body.[253]

Cancer is caused by genetic mutations. The high proofreading abilities of the cell's replication mechanism generally prevent replication errors, however rare mistakes do happen. There is approximately one mistake per $10^7$ nucleotides copied, and 99% of these replication errors are corrected by DNA mismatch. This increases the overall accuracy to one mistake per $10^9$ nucleotides copied. When copying mistakes are made, a miss-paired nucleotide is left behind, and if this is not corrected it will result in a permanent mutation in the next round of DNA replication. Mismatch repair proteins see the error, remove that portion of the DNA strand, and then resynthesize the missing DNA. This repair mechanism restores the correct sequence. People who inherit one damaged mismatch repair gene are unaffected until the undamaged copy of the same gene is randomly mutated in a somatic cell. This mutant cell is then lacking in mismatch repair and will then accumulate mutations more rapidly than a normal cell. This cell then has a greater chance of becoming cancerous. Mismatch repair plays an important role in preventing cancer.[252]
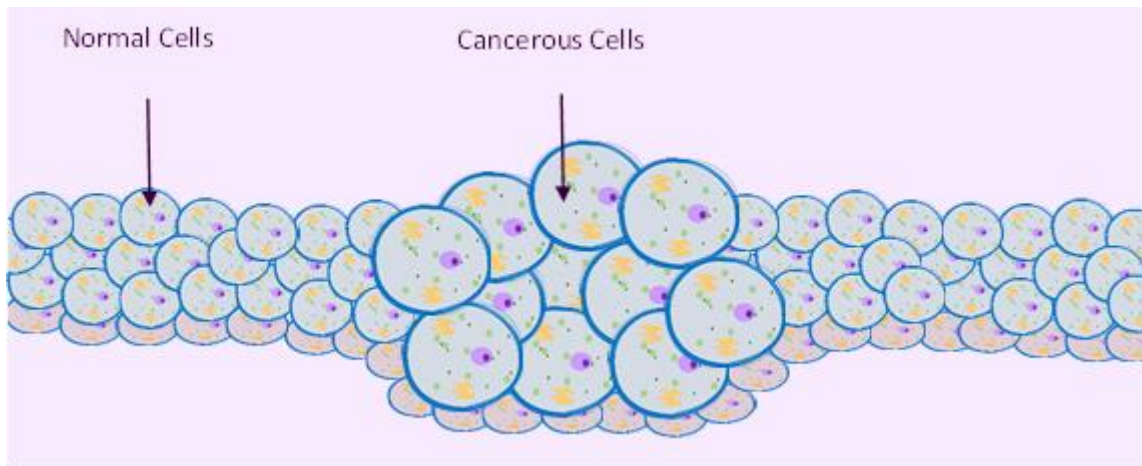
Figure 1.27. Cancerous cells

### 1.4.1.1 The Hallmarks of Cancer

Six hallmarks of cancer were proposed by Hanahan and Weinberg in 2000.[254] They discussed the notion that as normal cells evolve progressively to a neoplastic state they acquire a succession of these hallmarks. Human tumour pathogenesis is a multistep process which by emerging cancer cells begin to acquire traits that allow them to become tumorigenic and malignant.[255] Over the decade following, this research has developed and extended to reveal that the biology of tumours is more than that traits shown by the cancer cells but also must incorporate that contributions of the tumour microenvironment to tumorigenesis. The original six hallmarks include sustaining proliferative signalling, evading growth suppressors, resisting cell death, enabling replicative immortality, inducing angiogenesis, and activating invasion and metastasis. This list has been enhanced to include a more up to date picture of cancer which can see seen in figure 1.28.
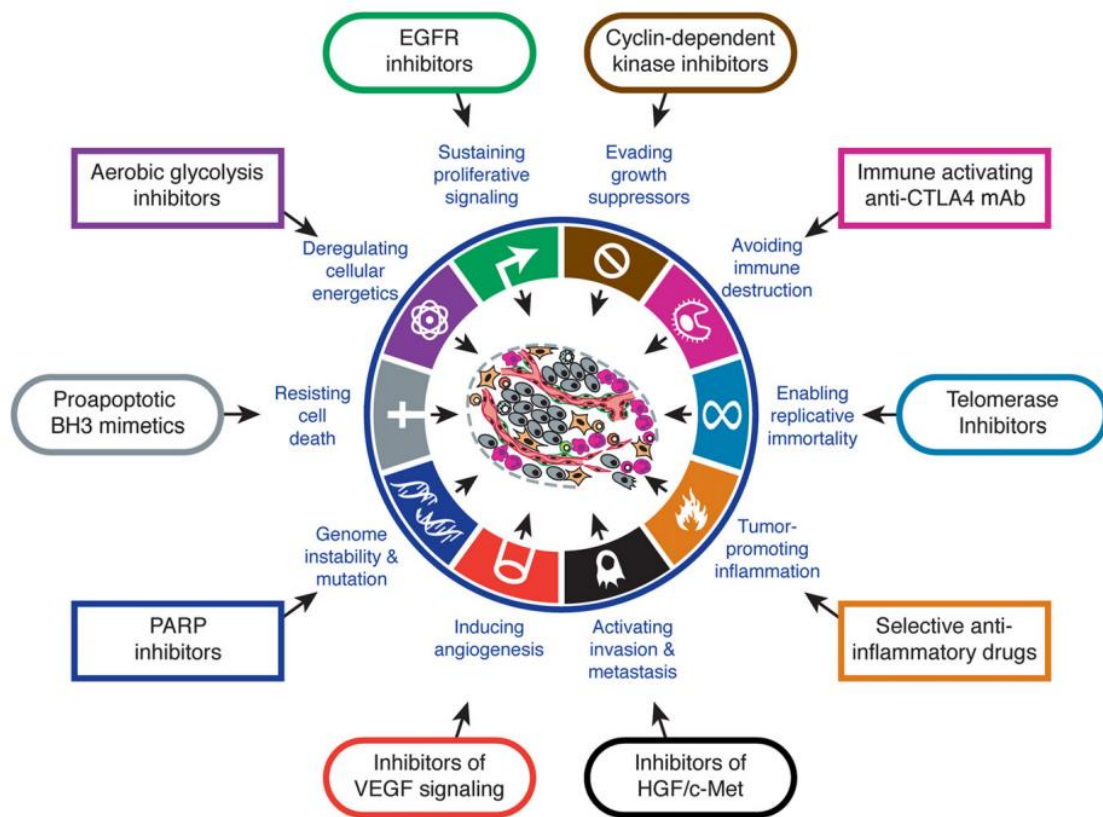
Figure 1.28. Therapeutic Targeting of the Hallmarks of Cancer Drugs that interfere with each of the acquired capabilities necessary for tumour growth and progression.[255]

The progression in the idea of the hallmarks of cancer has produced suggested therapeutic targets such as EGFR inhibitors (figure 1.28).

# 1.5. Epidermal Growth Factor Receptor and Cancer

## 1.5.1. The Epidermal Growth Factor Receptor
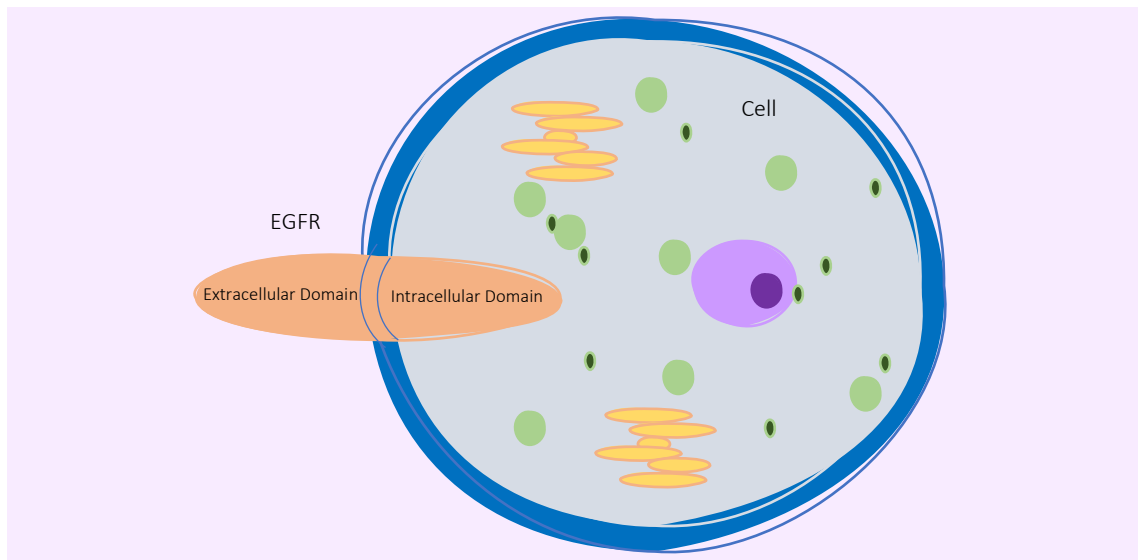


Figure 1.29. EGFR Transmembrane Protein.

EGFR is a transmembrane protein that is a receptor for members of the epidermal growth factor family of extra cellular protein ligands (170 kDa) (figure 1.29).[251] It is a member of the ErbB family of receptors which is a subfamily of four closely related receptor tyrosine kinases.[256] EGFR an important role in determining structure and 3D conformation of the EGFR protein.[257] Kinase is activated by the binding of EGF (figure 1.30) to the receptor and this proceeds to receptor autophosphorylation. The autophosphorylation is crucial for the interaction of the receptors and its substrates. EGF binds to the receptors by the SH2 domain.[258]



Figure 1.30. The Crystal Structure of EGF Protein.[259]

EGF together with EGFR also play a vital role in wound healing through stimulating epidermal and dermal regeneration.[260] EGFR signalling can also regulate several functions such as differentiation, motility, apoptosis and cell proliferation.[261] EGFR and members of its family are the key contributors of a multifaceted signalling cascade that controls growth, migration, differentiation, signalling, adhesion, and survival of cancer cells. Owing to their multidimensional role in the progression of cancer, EGFR and its family members have emerged as attractive candidates for anti-cancer therapy.[262] Specifically, the aberrant activity of EGFR has been shown to play a key role in the development of growth and tumour cells, where it is involved in numerous cellular responses including proliferation and apoptosis.[263]

### 1.5.2.1. The ErbB Family

The ErbB family of receptor comprises four known members namely ErbB/EGFR/human epidermal growth factor receptor (HER1), ErbB2/HER2/Neu, ErbB3/HER3 and ErbB/HER4. They are all transmembrane receptor tyrosine kinases with molecular weights ranging from 170 to 185 KDa.[264] Structurally, the ErbB family members consist of a cysteine-rich extracellular N-terminal ligand binding domain and a dimerization arm, a hydrophobic transmembrane domain and an intracellular, highly conserved, cytoplasmic C-terminal tyrosine kinase domain with several tyrosine residues that are the target of autophosphorylation. The extracellular region of EGFR is subdivided into four domains (I, II, III and IV) (figure 1.31).[265,266]
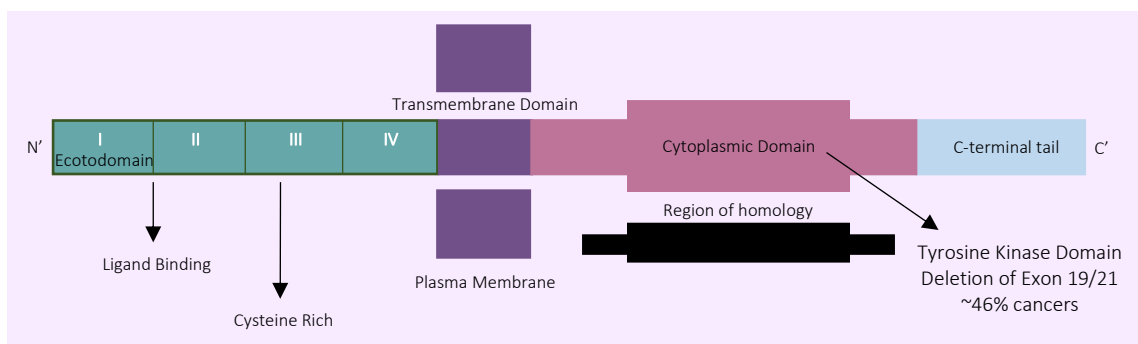


Figure 1.31. Schematic of EGFR receptor tyrosine kinase.

The crystal structure of the ectodomains of EGFR, HER3 and HER4 reveal two distinct conformations:

1. A closed inactive conformation. In the closed conformations, domains II and IV interact with each other at the intermolecular level, thus preventing domains I and III from interacting with their cognate ligand.[267, 268]

2. An open, active conformation. The open conformation is facilitated by the moving away of domains II and IV, thereby enabling domains I and III to expose their ligand-binding pocket and react with their corresponding ligand. As a result the dimerization arm in domain II then interacts with an identical dimerization arm of another receptor molecule to form a homodimer.[265,266]

Both the open and closed conformation remain in equilibrium with each other.[269] However, the binding of a ligand shifts the equilibrium and stabilises the open conformation, further enabling the accumulation of active homodimers and maintaining active receptor signalling. EGFR promotes heterodimerization with other members of the HER family including HER1, HER3 and HER4. EGFR may initiate cellular signalling cascades by itself through homodimerisation or through transactivation with other HER family members. EGF can induce heterodimerization of EGFR with HER2, HER3 or HER4.[270] ErbB family members can be activated by 13 known ligands which include EGF, transforming growth factor receptor alpha (TGF-α), amphiregulin (AR), betacellulin (BTC), heparin-binging EGF-like growth factor (HB-EGF), epiregulin (EPR), epigen (EPG) and neuregulin's 1-6 (NRG).[270, 271]

### 1.5.2.2. Tyrosine Phosphorylation and De-phosphorylation

Protein tyrosine kinases catalyse the transfer of the y-phosphate from ATP to the hydroxyl group of tyrosine residues. Protein tyrosine phosphatases remove the phosphate group from phosphotyrosine (figure 1.32). EGFR binds to its cognate ligand EGF, which further induces

tyrosine phosphorylation and receptors dimerization with other family members leading to enhanced uncontrolled proliferation.[272] Activation of EGFR signalling is triggered by ligand-induced receptor dimerization. Tyrosine residues present in the intrinsic kinase domain of one receptor will then cross phosphorylate specific residues in the C-terminal tail of the partnering receptor. This provides a scaffold for the recruitment of effector proteins.[273, 274] This happens by the Src homology 2 (SH2) and phosphotyrosine binding domain on the effector proteins and the phosphotyrosine motif on the intracellular tyrosine kinase domain of the receptor. After dissociation, the activated adaptor and effector proteins will then stimulate their corresponding signalling cascades. This pathway includes the KRAS-BRAF-MEK-ERK, phosphoinositide 3-kinase (PI3K), phospholipase C gamma protein pathway, and anti-apoptotic AKT kinase pathway and the STAT signalling pathway. This all leads to cell migration, survival, proliferation, adhesion and angiogenesis.[270, 275] These cell processes are generally deregulated in malignant cells. This is because of the several mutations concealed in various genes are also integrated in these pathways.
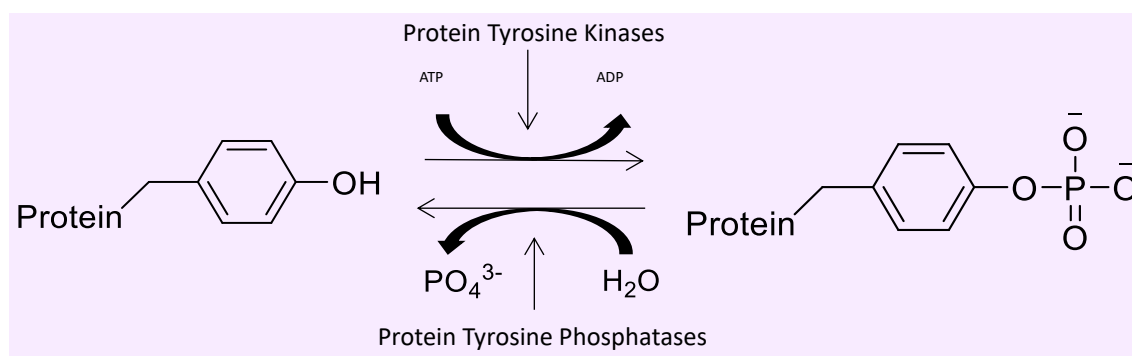


Figure 1.32. Scheme of Tyrosine Phosphorylation and Dephosphorylation.

### 1.5.2.3. EGFR Signaling Pathway

EGFR is a particularly important protein as it is involved with and activates many signalling cascades (figure 1.33). The ligand induced receptor dimerization and the successive autophosphorylation of tyrosine residues creates docking sites for numerous membrane

targeted proteins.[276] The cytoplasmic mediators that bind to EGFR phosphotyrosine residues are either adaptor proteins or enzymes such as the sarcoma viral oncogene human homolog (c-Src) and phospholipase C gamma 1 (PLC). As seen in Figure 1.30, examples of adaptor proteins are SHC transforming protein 1 (SHC), growth factor receptor-bound protein (GRB2), Cas-Br-M ecotropic retroviral transforming sequence (c-Cbl) and NCK adaptor protein 1 (NCK). The adaptor proteins SHC and GRB2 convert the guanine nucleotide exchange factor son of sevenlees homolog 1 (SOS) and form the complex SHC-GRB2-SOS. Once the SOS is activated, it in turn activates small GTPase v-Ha-ras Harvey rat sarcoma viral oncogene homolog (Ras) by its change from the inactive GDP-binding state to the active GTP-binding state. Activated Ras stimulates v-Raf-1 murine leukemia viral oncogene homolog 1 (Raf), mitogen-activated protein kinase kinase 1 and 2 (MEKK1 and MEKK2), mitogen-activated protein kinase 1 and 3 (ERK1/2) kinase cascade. This results in activation of the transcription factors ELK1, a member of ETS oncogene family and v-Myc myelocytomatosis viral oncogene homolog (c-Myc) and v-Fos FBJ murine osteosarcoma viral oncogene homolog (c-Fos).[277] The adaptor NCK couples EGFR stimulation to the activation of MAPL cascade called the JNK kinases cascade NCK1 then recruits p21-actviated kinases 1 (PAK1). The NCK/PAK1 complex binds Mitogen-activated protein kinase kinase kinase 10 (MLK2) which then activates the JNK cascade with consists of MLK2, MKK4/7 and Mitogen-activated protein kinase 8 and 9 (JNK1 and JNK2). The employment of this cascade to the activated membrane receptor then localises MLK2 on the plasma membrane where it is activated by known upstream effectors such as Rac1. The stimulation of JNK cascade allows the activation of the transcription factors Elk1 and c-Jun.[278] Phosphatases 1 and 4 (MKP-1 and MKP-2) reduce the effect of the JNK cascade.[279] GRB2 binds with its SH3 domain to the proline-rich regions of the c-Cbl protein. c-Cbl is a tyrosine phosphorylated by tyrosine kinases when stimulated with EGFR. The EGF stimulation prompts the connection between c-Cbl with the regulatory p85 subunit of phosphatidylinositol 3 kinase (PI3K reg class IA (p85)).[280] Once PI3K reg class IA is activated it converts Phosphatidylinositol 4,5-biphosphate (PtdIns(4,5)P2) to

Phosphatidylinositol 3,4,5-triphosphate (PtdIns(3,4,5)P3).[281] PtdIns(3,4,5)P3 works with the inner surface of the plasma membrane. It stimulates the recruitment of proteins with pleckstrin homology domains. One of these proteins is the serine/threonine kinase AKT. This protein is an essential mediator of important cell processes, such as apoptosis, cell cycle, protein synthesis and the regulation of metabolism.[282] EGFR recruits the phosphorylase PLC-gamma 1.[270] This phosphorylated PLC-gamma 1 then produces Diacylglycerol (DAG) and Inositol-1,4,5-trisphosphate form PtdIns(4,5)P2.[283] DAG activates many forms of Protein Kinase C (PKC) such as alpha, beta and gamma, which then phosphorylate and activate c-Raf-a which leads to amplification of kinase cascades Ras/MEK1 and MEK2/ERK1/2.[284,285] PKC family members also activate the Nuclear factor NF-kappa-B inhibitor kinase (IKK) which then activates the Nuclear factor NF-kappa-B (NF-kB) inside the nucleus.[286] The cytoplasmic tyrosine kinase c-Src, which is employed in many important cellular processes, for example cytoskeletal organisation and mitogenic signalling. c-Src can be stimulated by EGF locking to EGFR and then stimulates Shc, Signal transducer and activator of transcription 3 (STAT3), cytoskeletal components and some other proteins.[287] EGFR will induce cell migration via janus kinase 1 and 2 (JAK1/2) with STAT1 and STAT3.[287] EGF moderates the epithelial-to-mesenchymal transition (EMT). Too much stimulation of EGF can lead to EMT during tumour development.[288, 289] Generally, EGF increases Transforming growth factor beta (TGF-beta) signalling and encourages EMT only when combined with TGF-beta 1.[290,291,292] EGF dependent EMT commonly uses the pathways ERK1/2[290, 293] and PI3K activation.[291] EGF activates Integrin-linked kinase (ILK) signalling Glycogen synthase kinase 3 beta (GSK3 beta) and AKT-dependent activation of Cyclin E and Cyclin-dependent kinase 2 (Cdk2) via PI3K, which promotes proliferation during EMT.
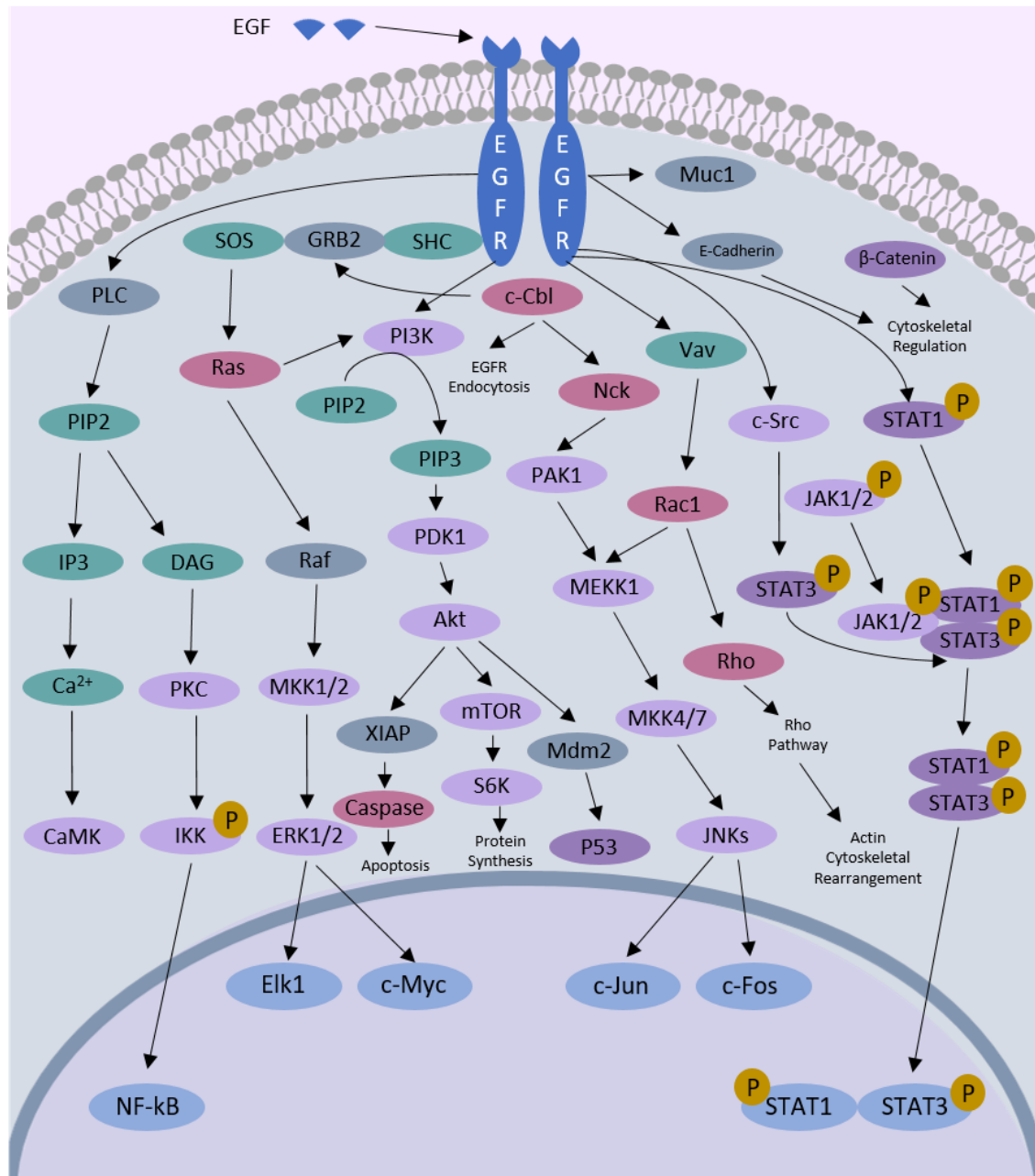
Figure 1.33.  EGFR Signaling Pathway.

### 1.5.2.4. Why is EGFR a good drug target?

Several lines of evidence have identified EGFR as a good drug target for anticancer therapy.[294] It is connected to the control of cell survival, metastasis, proliferation and angiogenesis[295] as seen by the signalling pathways. EGFR is a good target because it often is expressed at a high level in many solid tumours such as neck, breast, and head, and is also seen in prostate cancer. In some tumours deregulation of EGFR function has been associated with tumours in an

advanced stage[296] and resistance to hormone therapy, chemotherapy and radiation. In over 5000 patients with breast cancer, the mean percentage of EGFR positivity was 45% with a range of 14-91%.[297]

These frequent somatic mutations and overexpression of EGFR in cancerous cells provides a therapeutic window for targeting the receptor.[3] The oncogenic pathways driven by EGFR are in a complex that regulate the activity of pathway components in response to interactions. Because of this there are many potential targets within these pathways, there is the potential that direct inhibition of one protein will affect the activity of another protein and pathway and neutralise the growth suppression caused by the inhibition of the targeted pathway.[3] As an example c-Raf in Ras-MAPK pathway is negatively regulated by activated AKT which causes PI3K/ARK inhibitors to drive activation of MAPK signalling.[298,299] Therefore, individual targeting of the downstream effectors of EGFR is challenging and may need a combinatorial approach which targets several areas in the pathways. EGFR is located at the peak of this complex signalling network; this is an advantage for direct targeting of the functions of the receptor instead of other components of the pathways it activates. Moreover, as it is a cell-surface receptor and a druggable kinase that is over-expressed and genetically altered in cancer cells, it makes EGFR is a very attractive therapeutic target for cancer therapy.

### 1.5.2.5. EGFR Targeted Therapies

EGFR is functionally involved in diverse cellular processes. Several approaches have been developed that target and interfere with the EGFR mediated effects. Two different therapeutic approaches currently employed for targeting EGFR in various human malignancies are the use of monoclonal antibodies and small molecule tyrosine kinase inhibitors. The distinct mechanism of actions is that anti-EGFR antibodies bind to extracellular domains and tyrosine kinase inhibitors that target the intracellular tyrosine kinase domain.[300]

### 1.5.2.5.1. Anti-EGFR monoclonal antibodies

The monoclonal antibodies against EGFR were specifically designed to be directed against the extracellular region of EGFR. This creates a ligand competitive inhibition which then prevents auto-phosphorylation receptor dimerization, and downstream signalling.[301] As well as inhibiting EGFR signalling, these monoclonal antibodies induce receptor internalization, ubiquitination, degradation and prolonged downregulation.[302] There are two other proposed mechanisms, where the binding of the monoclonal antibody can lead to the induction of antibody dependent cell mediated cytotoxicity, which results in endocytosis.[303] Cetuximab is an example of a monoclonal antibody against EGFR. Cetuximab is a human-murine chimeric anti-EGFR monoclonal antibody approved by the US Food and Drug Administration (FDA) for treatment of advanced bowel cancer, head and neck cancers and the start of mouth and throat cancer (figure 1.34).[304] Cetuximab binding to domain III of EGFR blocks downstream signalling by stimulating receptor internalisation and hindering ligand receptor communication. Of people who have KRAS wild type tumours in colorectal cancers, 60% of them respond successfully to Cetuximab treatment.[305] It is generally used in patients who have failed with other types of chemotherapy such as irinotecan and oxaliplatin.[306] The tumour regression rate was 10.8% in patients using cetuximab with the tumour growth delayed by 1.5 months.[307] A recent study trialled a combined therapy of cetuximab with oxaliplatin and capecitabine, however this did not show any benefit to patients with wild type KRAS tumours.[308, 309] In 2004 the FDA approved cetuximab combined with a platinum-based therapy for the treatment of patients with squamous cell carcinoma of the head and neck with metastatic disease and alongside radiation therapy for locally advanced cancer[310]. In 2008 the Committee for Medical Products for Human Use (European Medicines Agency (EMA)) approved cetuximab for patients with advanced colorectal cancer who had 75% EGFR positive expression and wild type K-Ras in their tissues and who been unsuccessful with oxaliplatin or irinotecan chemotherapy.[311] In 2021 the FDA approved a new dosage regimen for

cetuximab of 500mg/m$^2$ as a 120 minute intravenous infusion every two weeks for patients with K-Ras wild type, EGFR expressing colorectal cancer or squamous cell carcinoma of the head and neck.[312]



Figure 1.34. Protein Crystal Structure of Cetuximab.[313]

### 1.5.2.5.2. Small molecule tyrosine kinase inhibitors

Small molecule inhibitors (SMIs) are some of the most effective drugs for targeted cancer therapy. There is an increasing amount of approved SMIs of receptor tyrosine kinases (RTKs) such as tyrosine kinase inhibitors (TKIs) in clinical oncology. This increase shows that this becoming a popular therapeutic tool.[314] Gefitinib (figure 1.35)[315] approved in 2003 and Erlotinib[316] approved in 2004 are some of the first TKIs created for the selection of EGFR-TKIs for non-small cell lung cancer patients.[317,318] Erlotinib in 2005 was also approved for the treatment of patient with metastatic pancreatic cancer when partnered with Gemcitabine.[319] The anti-tumour effects of both drugs have been researched in other cancer types such as cervical[320], esophageal[321], gastric[322] hepatocellular carcinoma[323] and renal cell carcinoma.[324] Erlotinib has been shown to be successful as a first line treatment in gastroesophageal cancer, however is has no clinical benefits in gastric cancer.[322] These drugs are first generation of EGFR inhibitors. In 2022, research is now moving forward on to fourth generation EGFR inhibitors.
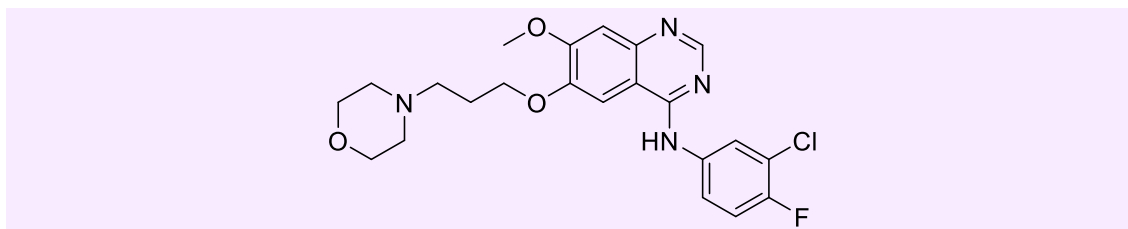
Figure 1.35. Chemical Structure of Gefitinib.

### 1.5.2.5.3. The progression of EGFR inhibitors

EGFR tyrosine kinase inhibitors have revolutionised the treatment of EGFR-mutant non-small cell lung cancer (NSCLC); however, secondary resistance limits their efficacy, emphasizing the need for newer approaches.[325] TKIs against EGFR are now standard treatment in the clinic for patient with advanced EGFR mutant NSCLC.[326]

First-generation EGFR TKIs, of which bind competitively and reversibly to the ATP-binding site of the EGFR tyrosine kinase domain, have produced significant improvement in the outcomes for NSCLC patients with activating EGFR mutations L858R and Del19.[327] However, after 12 months of treatment, all patients developed tumour resistance. In over half of these patients, this was due to the development of the EGFR T790M resistance mutation.[328]

The second-generation EGFR/HER TKIs were developed to help treat this resistant disease, targeting T790M and EGFR-activating mutations and wild-type EGFR.[329] They exhibited promising anti-T790M activity *in vitro* and *in vivo*, however failed in clinical trails because of dose-limiting toxicity due to simultaneous inhibition of wild-type EGFR.[330]

The third-generation of EGFR TKIs target EGFR T790M selectively and irreversibly, showing promising effectiveness in NSCLC resistant to the first- and second-generation EGFR TKIs.[331] These EGFR TKIs also appear to have lower incidences of toxicity due to the limited inhibitory effect on wild-type EGFR. Currently, Afatinib has been approved for first-line treatment of NSCLC with activating EGFR mutations.[332]

In the third-generation EGFR TKIs, Osimertinib is currently the only drug approved by the FDA the EMA for treatment of metastatic EGFR T790M NSCLC patients who have progressed on or after EGFR TKI therapy.[333] EGFR TKIs research is still ongoing, paving the way for fourth-generation TKIs, leaving space for more EGFR targeted research and development.

### 1.5.2.5.4. Current Anti EGFR Aptamers

Aptamers have been selected for EGFR, such as J18 aptamer ($K_d$ of 7 nM),[334] E03, E04 and E07 aptamers, having E07 as the most successful, binding tightly to the wild type receptor with a $K_d$ of 2.4 nM (figure 1.36). E07 aptamer has been shown to compete with EGF for binding to EGFR. It also binds to a novel epitope of EGFR and a deletion mutant EGFRvIII. EGFRvIII is commonly found in breast, lung and grade IV glioblastoma multiforme (which is currently unresponsive to other therapies). E07 aptamer binds to cells that express EGFR, it blocks the receptor autophosphorylation and also prevents proliferation of tumour cells.[6]



| $K_d$ (nM) | Aptamer E03 | Aptamer E04 | Aptamer E07 |
|---|---|---|---|
| hEGFR | $3.0 \pm 0.3$ | $2.6 \pm 0.3$ | $2.4 \pm 0.2$ |
| mEGFR | $29 \pm 1$ | $48 \pm 3$ | $36 \pm 4$ |
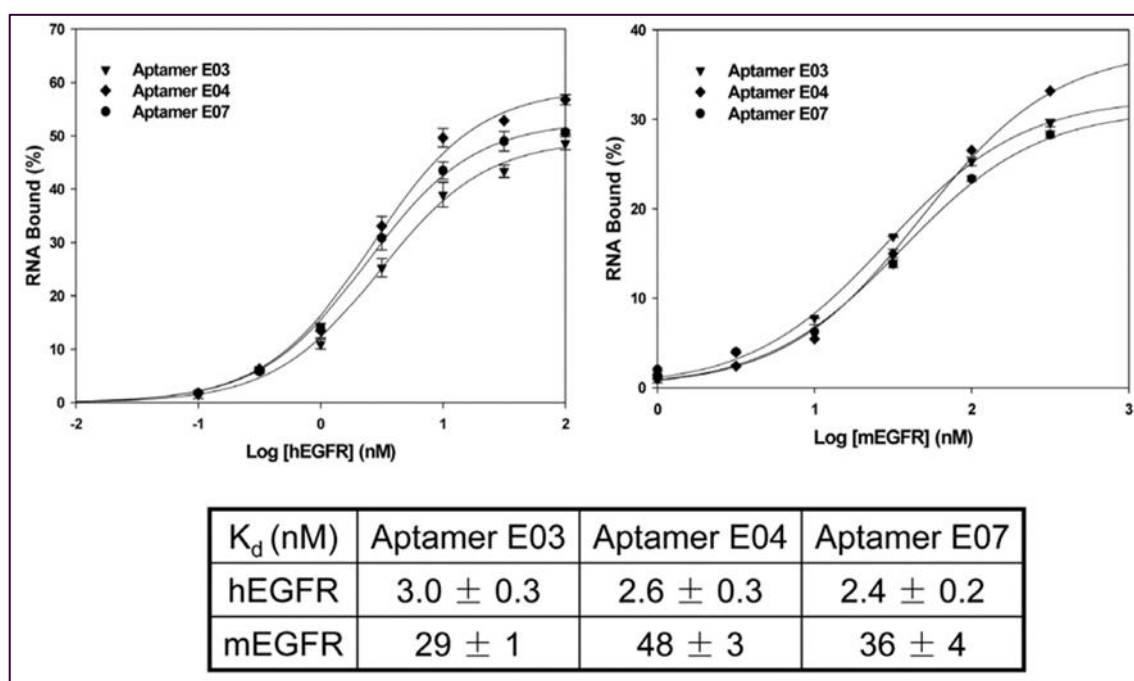
Figure 1.36. Binding constants for Aptamers E03, E04, and E07. Binding was done using 0.1 nM aptamer and varying amounts of human EGFR or mouse EGFR. Binding assays were carried out in triplicate and the average value and standard deviation are shown.[6]

E07 was then remodelled into a short aptamer, MinE07. MinE07 still retained the good binding affinity for EGFR. MinE07 and Cetuximab were run on a cell growth assay, and the results showed that a much larger dose of MinE07 was needed to achieve the same results a Cetuximab. This aptamer is a very promising therapeutic candidate or drug delivery vehicle, however a lot more research needs to be done to achieve its full potential, which produces a need for this research project.

## 1.6. Project Aims

It is vital to have EGFR-targeted therapies because it is among the most commonly altered genes in cancer.[3] EGFR promotes oncogenesis in many cancers, such as breast, lung, colon and glioblastoma.[1]

The recognition capacity of aptamers can be harnessed for therapeutic agents. Aptamers chemistry is mainly limited to that of nucleic acids, and although non-natural modifications of nucleic acids are known to enhance aptamer affinity,[5] there is not yet a technology or method for selecting the right modifications amongst billions of possibilities.

This project aims to develop the first general method for the discovery of nucleoside modifications that increase aptamer binding efficacy. A library will be created of over 65,000 different chemical modifications on a known aptamer sequence (MinE07)[6], which binds to EGFR protein. The modifications will use different types of chemistry to see how they will affect the binding and folding of the aptamer MinE07, aiming to improve the binding between MinE07 and EGFR, these modifications will be different from the standard nucleobases found in RNA. This library will be attached on beads such that each bead displays millions of copies of a single aptamer sequence. The selection will then be carried out using a flow cytometer (FACS) to separate out the aptamer sequences with the highest affinity for EGFR from the one-bead-one-sequence aptamer library. These will then be identified using MS/MS technology. These top

selected aptamer sequences will be subjected to biophysical and biological testing. This project has potential impacts for: a new method for synthesising and screening large drug candidate libraries, accelerating the drug discovery processes and proof of concept potentially providing new aptamer drug candidates or delivery vehicle targeting EGFR protein.

## 1.6.1. Project Objectives

This research project has six main objectives shown below to achieve the project aims of creating the first general method for the discovery of nucleoside modifications that increase aptamer binding efficacy as well as proving a now potential EGFR targeting therapy.

1. Selection technique development involving flow cytometry and one-bead-one-compound methods

2. Nucleic acid modification design and synthesis

3. Synthesis of aptamer library

4. Selection of top aptamer sequences using flow cytometry

5. Identification of top selected aptamers using MS/MS technology

6. Biophysical and biological testing on selected aptamer sequences

# Chapter 2

## Selection Technique Development involving Flow Cytometry and the One-Bead-One-Compound Method

# Chapter 2

## 2.1. Introduction

### 2.1.1. Solid-Phase Synthesis

Solid-phase synthesis is a method where molecules are covalently bound on to a solid support, allowing them to be built step by step followed by cleavage from the solid support (figure 2.1). There is a wide range of materials being used as solid supports. The synthesis is generally step-by-step in a single container applying protection group chemistry. This method produces high yields by using excess reagent. Compared with liquid state synthesis it has a number of benefits such as high efficiency and throughput, increased speed and simplicity.[335] Impurities and excess reagents are washed away so no further purification is needed after. The process can also be automated on computer-controlled synthesisers. This is useful in drug discovery processes to synthesise small organic molecules on to a solid support to produce high yields.[336]

Solid-phase synthesis has been used for the synthesis of DNA, RNA, peptides[335, 337] and other small molecules. This method is now used in combinatorial chemistry. It was developed first in 1950's by Robert Bruce Merrifield for synthesis of peptide chains,[338, 339] winning him a Nobel prize in 1984.[340]



Solid support                    Linker Addition                    Functional
                                                                     Group Addition
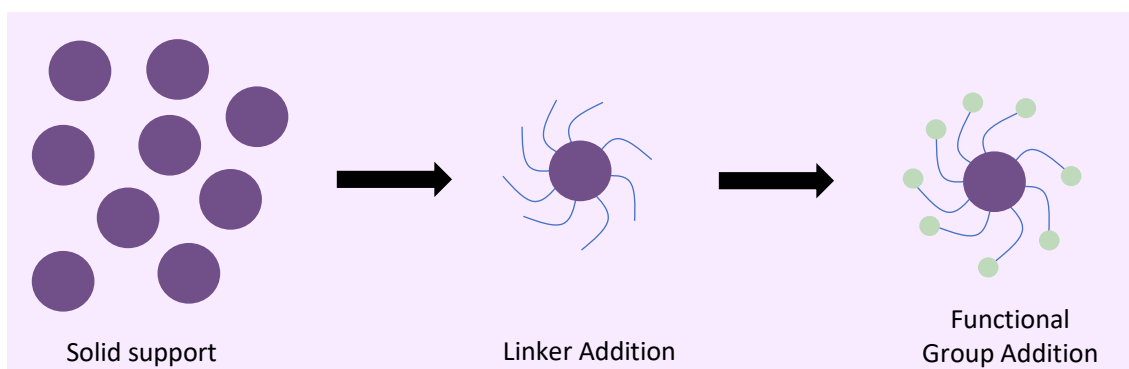
Figure 2.1. Solid-Phase Synthesis

### 2.1.1.1. Solid Supports

Solid supports are insoluble particles. They have a typical diameter of 10-200 μm, which the compounds can bind to during the synthesis. There are many variations of solid support that have been used, such as controlled pore glass and polystyrene.

### 2.1.1.1.1. Controlled Pore Glass

Controlled pore glass (CPG) is non-swelling and rigid. It has deep pores for the synthesis of molecular chains within the beads, greatly increasing the surface area. 500 Å pore glass supports are used for the synthesis of short oligonucleotides. Larger pore glass supports are more fragile, however 1000 Å can be used for oligonucleotides up to 100 bases.[341] It is well-established that CPG is a suitable support for automated DNA synthesis,[342] because this material is chemically inert and is stable in the presence of solvents and reagents during synthesis. Its porosity offers a large surface area, which allows high loading capacity of functional groups or biomolecules.[343]

### 2.1.1.1.2. Polystyrene

Polystyrene beads that are highly cross-linked work well for oligonucleotide and peptide synthesis because they have good moisture exclusion properties. Satisfying yields can be obtained using the copolymer of polystyrene and polyethylene glycol, commercially known as Tentagel®, in stereoselective solid-phase glycosylations.[344] For oligonucleotide synthesis, solid supports are typically made with a loading of 20-30 μmol of nucleoside per gram of resin.[345] At higher loading, oligonucleotide synthesis becomes less efficient because of steric hindrance between adjacent DNA chains attached to the resin.[346] However, polystyrene supports with loadings of up to 350 μmol/g can be used in some applications, such as short oligonucleotide synthesis, and it enables the synthesis of large quantities of oligonucleotides.[347]

### 2.1.1.1.3. Disadvantages of Solid Supports

Although a powerful technique, solid-phase synthesis also has some limitations. The primary limitation being the need for high coupling yields at every step in the synthesis. This is because the overall yield rapidly decreases, the more steps there are. Such that the yield of a trinucleotide chain with 90% yield at each step is 73% (*0.9 x 0.9 x 0.9 x 100% = 73%)*. The consequence of the exponential relationship between overall yield and average coupling efficiency is that long oligonucleotides can only be prepared to a good standard with every step producing a high yield, and this is the major reason why solid-phase oligonucleotide synthesis was not successful until the early 1980s.[348] High coupling efficiencies are now achieved and have allowed for oligonucleotides to be synthesised as long as 200 bases.

## 2.1.2. Chapter Aims

The aim of this chapter is to explain the choice of solid support for the synthesis of the aptamer library by showing it is suited for organic synthesis, is homogenous in size and compatible for sorting in the FACS and has the synthetic capacity to hold a concentration high enough to be observed by mass spectrometry. The homogeny of size will be shown visually by Scanning Electron Microscopy (SEM) and optical microscope. Its suitability for use in organic synthesis will be demonstrated by linkers, fluorescent tags being bound successfully to the solid support followed by synthesis of the parent aptamer. The fluorescently tagged solid supports will be run through the FACS to check the stability of the solid support and also the stringency of the FACS picking out different levels of fluorescence.

## 2.2. Materials and Instrumentation

### 2.2.1. Materials

TentaGel® M NH$_2$ Monosized Amino TentaGel Microspheres were purchased from Rapp Polymere. Sheath Fluid was purchased from BDBiosciences. Micro glass spheres were purchased from Sigma Aldrich. [3-(2-aminoethylamino)propyl]trimethoxysilane], 6-carboxyfluroescein, 10-hydroxydecanoic acid, boric Acid, DCM, DMF, EDTA, ethanol, KCN, nbutanol, N,N'-Dissopropylcarbodiimide, ninhydrin, rhodamine B, PBS, phenol, pyridine, tris base were purchased from Thermo Fisher Scientific. Synthesiser reagents: oxidizer (0.02M iodine, 20% pyridine), Cap A Mix (THF/Pyridine/acetic anhydride 8:1:1), Cap B Mix (10% methylimidazole in THF), deblock (3% trimethylamine in DCM) and ETT activator solution (0.25 M, 5-ethylthio-1H-tetrazole in acetonitrile) were purchased from Link Technologies. N2-acetyl-2'-O-tert-butyldimethylsilyl-5'-O-DMT-guanosine 3'-CE phosphoramidite, N6-benzoyl-2'-O-tert-butyldimethylsilyl-5'-O-DMT-adenosine 3'-CE phosphoramidite, N4-acetyl-2'-deoxy-5'-O-DMT-2'-fluorocytidine 3'-CE phosphoramidite and 2'-deoxy-5'-O-DMT-2'-fluorouridine 3'-CE phosphoramidite were purchased from CarboSynth.

#### 2.2.1.1. Buffers

10 x tris-borate buffer (TBE): tris Base (108 g), boric Acid (55 g), EDTA (7.45 g), dH$_2$O to 1 L at 8.3 pH.

10 x TAMg buffer (TAMg): tris base (54.51 g), Mg(OAc)$_2$.4H$_2$O (26.8 g) dH$_2$O to 1 L at 7.8 pH.

## 2.2.2. Instrumentation

### 2.2.2.1. Expedite™ 8909 DNA Synthesiser

All library components were synthesised on an Expedite™ 8909 Nucleic Acid Synthesiser system provided by Biolytic. Phosphoramidites were dissolved in dry acetonitrile to the concentrations as suggested by the supplier, solvents were used as provided. Oligomers were synthesised on a 1 µM scale.

### 2.2.2.2. Fluorescence Activated Cell Sorter

Flow cytometry data was collected on a BD FACSJazz™ Cell Sorter by Becton Dickinson. All flow cytometry data was processed using BD FACS™ Sortware sorter software.

### 2.2.2.3. Fluorescence Microscope

Leica DMR microscope equipped with a Leica DFC9000 GT digital camera. The illumination source was a CoolLED pE-300[ultra] fluorescence microscopy Illumination System. Images were acquired using Leica Application Suite X software. Excitation 495nm and emission 525/50 nm. Excitation 515-560 nm and emission long pass 560-590 nm.

### 2.2.2.4. Scanning Electron Microscopy

SEM images were obtained on a Hitachi S-3400N, using a 10 kV electron beam and secondary electron detector.

### 2.2.2.5. UV-Visible Absorption Spectra

UV- visible absorption spectra were recorded on a NanoDrop One UV-Vis spectrophotometer by Thermo Fisher Scientific.

## 2.3. Experimental

### 2.3.1. Imaging of micro glass spheres by optical microscopy

A small sample of micro glass spheres in water was smeared onto a microsphere glass slider. This was placed on the microscope platform and visualised and images were taken.

### 2.3.2. Imaging of micro glass spheres using SEM

SEM images were obtained on a Hitachi S3400N, using a 10 kV electron beam and secondary electron detector. The sample was deposited on carbon tabs placed on aluminium sample holders. The images were visualised on ImageJ and the data analysis was done using Origin Lab software.

### 2.3.3. Stability of micro glass spheres in solution

Samples of micro glass spheres were put into 4 different solvents; autoclaved water, which was reverse-osmosis purified to a resistance of 18.3 megohms/cm (AC water), tris-borate buffer (TBE), TAMg buffer (TAMg) and phosphate buffered saline (PBS), sonicated and left-over night (Table 2.1).

Table 2.1. Micro glass sphere samples in different buffers

| Amount of Micro glass spheres (g) | Buffer Used | Amount of Buffer (mL) |
|---|---|---|
| 0.1 | AC Water | 10 |
| 0.01 | AC Water | 10 |
| 0.001 | AC Water | 10 |
| 0.1 | TBE | 10 |
| 0.01 | TBE | 10 |
| 0.001 | TBE | 10 |
| 0.1 | TAMg | 10 |
| 0.01 | TAMg | 10 |
| 0.001 | TAMg Buffer | 10 |
| 0.1 | PBS Buffer | 10 |
| 0.01 | PBS Buffer | 10 |
| 0.001 | PBS Buffer | 10 |

**2.3.4. Kaiser Test**[349]

Reagent A: KCN (16.50 mg) was dissolved in distilled water (25.00 mL). 1.00 mL of this was then diluted with 49.00 mL of pyridine. Reagent B: Ninhydrin (1.00 g) in 20 mL of n-butanol. Reagent C: Phenol (40.00 g) in 20.00 mL of n-butanol. A few drops of reagent A, B and C were all added on to the sample. If a purple colour is seen, there are free primary amines present in the sample.

**2.3.5. Amination of micro glass beads**

Micro glass spheres (1.60 g) together with [3-(2-aminoethylamino)propyl]trimethoxysilane] (1.74 g) were suspended in ethanol (20.00 mL) and left to stir for 16 hours. The product was then collected by vacuum filtration and washed with ethanol (20.00 mL x 3). This produced a

white powdered solid (0.86 g, 53.4% ~11 – 22 μm sized beads), (0.17 g, 10.2%, ~2 – 10 μm sized beads yield). A Kaiser test was run on the product, showing a positive result.

### 2.3.6. Fluorescent tagging of the amine modified micro glass spheres

*N,N'*-Dissopropylcarbodiimide (0.63 g) was dissolved in dimethylformamide (20.00 mL). 6-carboxyfluroescein (1.08 g) and amine modified glass spheres (0.20 g) were added to the solution. The mixture was stirred for 3 hours. The product was centrifuged and the DMF supernatant removed. The product was washed with DMF (20.00 mL x 2) and with ethanol (20.00 mL x 2). The product was dried under vacuum filtration to produce an orange power (0.076 g, 38%).

### 2.3.7. 6-carboxyfluorescein micro glass spheres on the microscope

A small sample of 6-carboxyfluorescein tagged micro glass spheres in water were smeared onto a microsphere glass slide. This was placed on the microscope platform and visualised using a fluorescent lamp. Images of the sample was taken.

### 2.3.8. The Binding Capacity of the Micro Glass Spheres

Ten free amine samples of [3-(2-aminoethylamino)propyl]trimethoxysilane dissolved in water were made up as *per* Table 2.2. Ninhydrin (1.00 g) was dissolved in 20.00 mL of ethanol.

Table 2.2. Concentrations and amounts for the 10 free amine samples

| Concentration (%) | Water (mL) | Grams of Amine (g) |
|---|---|---|
| 100 | 25 | 0.0257 |
| 50 | 25 | 0.0129 |
| 20 | 25 | 0.00514 |
| 10 | 25 | 0.00257 |
| 5 | 25 | 0.00129 |
| 2 | 25 | 0.000514 |
| 1 | 25 | 0.000257 |
| 0.5 | 25 | 0.000129 |
| 0.2 | 25 | 0.0000514 |
| 0.1 | 25 | 0.0000257 |

100.00 μL of each amine solution was then added with 50.00 μL ninhydrin dye and vortexed. These samples were then measured on the Nanodrop UV-Vis spectrophotometer in triplicate and calibration graph was plotted. A 100.00 μL sample of amino tagged micro glass spheres was then added with 50.00 μL ninhydrin dye and vortexed. This sample was then run on the Nanodrop UV-vis spectrophotometer in triplicate.

### 2.3.9. Amidation of 10-hydroxydecanoic acid with TentaGel® M NH$_2$ Monosized Amino TentaGel Microspheres

*N,N'*-Dissopropylcarbodiimide (0.20 g) was dissolved in dimethylformamide (10.00 mL). 10-Hydroxydecanoic Acid (0.40 g) was added to the solution. TentaGel microsphers (0.10 g) were swelled in dimethylformamide (10.00 mL) and then added to the solution. This solution was stirred for three hours at room temperature. It was spun down and cleaned with

dimethylformamide (10.00 mL x 3) and then with ethanol (10.00 mL x 2). A white power was produced (0.95 g, 95%). A Kaiser test was conducted producing a yellow colour (successful).

## 2.3.10. Synthesis of MinE07ARP on TentaGel® M NH₂ Monosized Amino TentaGel Microspheres

MinE07ARP:5'-rGrGrAfCrGrGrAfUfUfUrArAfUfCrGfCfCrGfUrArGrArArArArGfCrAfUrGfUfCrArAr ArGfCfCrGrGrArAfCfCrGfUfCfC-3'.

The Expedite™ 8909 DNA synthesiser was used to synthesise the MinE07 parent aptamer (MinE07ARP) on the TentaGel microspheres. The G, A, C, and U phosphoramidites (1.00 g) being used were dissolved in 20.00 mL DCM and put in bottles, and then screwed into the synthesiser lines. The other reagents put on to the machine are oxidizer, cap A, cap B, deblock and activator. A leak test was run to check nitrogen is not leaking from the lines. If passed, the lines were then flushed with the new reagents added to them. –OH modified TentaGel microspheres (0.021 g) are added to the column, which is then placed on the synthesiser, which is then flushed with acetonitrile several times. The sequence and protocol are then selected using Expedite 8909 software connected to the Expedite™ 8909 DNA synthesiser. The aptamer sequences are then run and monitored using the trityl monitor. The protocol for coupling was extended 10x more than a standard DNA 1 µmol protocol.

### 2.3.11. Fluorescent tagging of TentaGel® M NH₂ Monosized Amino TentaGel Microspheres

#### 2.3.11.1. 6-Carboxyfluorescein

*N*,*N*'-Disopropylcarbodiimide (0.32 g) was dissolved in dimethylformamide (10.00 mL). The amine modified TentaGel microspheres (0.05 g) were swelled in dimethylformamide (10.00 mL), this was added to the solution along with 6-carboxyfluroescein (0.05 g). The solution was stirred for 3 hours at room temperature. It was spun down and cleaned with dimethylformamide (10.00 mL x 3) and then with ethanol (10.00 mL x 2). The product was then air dried and an

orange powered was collected (TGCFluor100) (0.074 g, 74%). Samples TGCFluor100, TGCFluor73 and TGCFluor36 are made per amounts in table 2.3.

Table 2.3 Amount of 6-Carboxyfluorscein for samples TGCFluor100/73/36. Calculations made by data given from TentaGel.

| Experiment | Percentage of 6-Carboxyfluorescein | Amount of 6-Carboxyfluorscein |
|---|---|---|
| TGCFluor 100 | 100% | 0.050 g (in excess, enough to cover $e^{19}$ binding locations on the microspheres) |
| TGCFluor 73 | 73% | 0.025 g (enough to cover $e^8$ binding locations on the microspheres) |
| TGCFluor 36 | 36% | 0.013 g (enough to cover $e^4$ binding locations on the microspheres) |

### 2.3.11.2. Rhodamine B

*N,N'*-Dissopropylcarbodiimide (0.15 g) was dissolved in dimethylformamide (10.00 mL). The amine modified TentaGel microspheres (0.05 g) were swelled in dimethylformamide (10.00 mL), this was added to the solution along with Rhodamine B (0.05 g). The solution was stirred for 3 hours at room temperature. It was spun down and cleaned with dimethylformamide (10.00 mL x 3) and then with ethanol (10.00 mL x 2). The product was then air dried and a pink powered was collected (TGRhodB100) (0.03 g, 60%). Samples TGRhodB100, TGRhodB73, TGRhodB36, TGRhodB17 an TGRhodB1.7 are made per amounts in table 2.4.

Table 2.4. Amount of Rhodamine B for samples TGRhod100/73/36/17/1.7. Calculations made by data given from TentaGel.

| Experiment | Percentage of Rhodamine B | Amount of Rhodamine B |
|---|---|---|
| TGRhodB100 | 100% | 0.050 g (in excess, enough to cover $e^{19}$ binding locations on the microspheres) |
| TGRhodB73 | 73% | 0.025 g (enough to cover $e^8$ binding locations on the microspheres) |
| TGRhodB36 | 36% | 0.013 g (enough to cover $e^4$ binding locations on the microspheres) |
| TGRhodB17 | 17% | 0.005 g (enough to cover $e^{1.9}$ binding locations on the microspheres) |
| TGRhodB1.7 | 1.7% | 0.0005 g (enough to cover $e^{0.19}$ binding locations on the microspheres) |

### 2.3.12. 6-carboxyfluorescein tagged TentaGel® M NH₂ Monosized Amino TentaGel Microspheres on the fluorescent microscope

A small sample of 6-carboxyfluorescein tagged TentaGel microspheres in water were smeared onto a microsphere glass slide. This was placed on the microscope platform and visualised using a fluorescent lamp. Images of the sample was taken.

### 2.3.13. Flow Cytometry analysis of tagged TentaGel® M NH₂ Monosized Amino TentaGel Microspheres

#### 2.3.13.1. 6-Carboxyfluorescein

The FACSJazz lasers were calibrated before the samples were run with the 8 peak BD beads. Plain microspheres were dispersed in 5 mL sheath fluid and put through the machine to check the calibration. Then samples TGCFluor100, TGCFluor73 and TGCFluor36 were ran through the FACS dispersed in 5 mL sheath fluid, being observed on laser 488 nm with filter 513/17 nm. All data was collected and analysed on BD FACS Sortware program.

#### 2.3.13.2. Rhodamine B

The FACSJazz lasers were calibrated before the samples were run with the 8 peak BD beads. Plain microspheres were dispersed in 5 mL sheath fluid and put through the machine to check the calibration. Then samples TGRhodB100, TGRhodB73, TGRhodB36, TGRhodB17 and TGRhodB1.7 were ran through the FACS dispersed in 5ml sheath fluid, being observed on laser 488 nm with filter 585/29 nm. All data was collected and analysed on BD FACS Sortware program.

#### 2.3.13.3. MinE07ARP

The FACSJazz™ lasers were calibrated before the samples were run with the 8 peak BD beads. Plain microspheres were dispersed in 5 mL of sheath fluid and put through the machine to check

the calibration. Then sample MinE07ARP was run through the FACS, dispersed in 5 mL sheath fluid, being observed on laser 488 nm with filter 585/29 nm. All data was collected and analysed using the BD FACS™ Sortware sorter software program.

# 2.4. Results and Discussion

## 2.4.1. Micro Glass Spheres

### 2.4.1.1. Characterisation of the micro glass spheres

The first beads analysed as a potential solid support candidate were micro glass spheres (9-11 µm). These were chosen to study first because of their success rate in being a solid support for oligonucleotide synthesis.[347,350] To characterise the size distribution of the micro glass spheres, they were analysed by optical microscopy and SEM. The images captured from the microscopy studies (figure 2.2) showed that the size range of the micro glass spheres was 2.4-21.5 µm instead of the stated 9-11 µm. Figure 2.2 shows that there were items in the samples that were not spheres. They may have been broken micro glass spheres, or some other foreign objects.



SEM of micro glass spheres            Micro glass beads under a Spheres
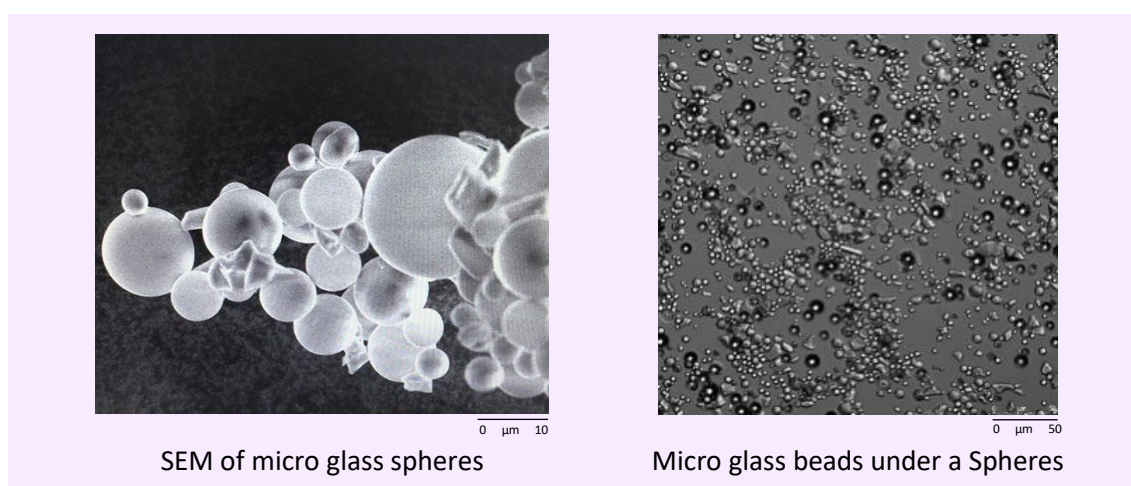
Figure 2.2. Microscopy images captured from SEM and Optical Microscope of the micro glass spheres.

A sample of micro glass spheres diameters were measured using SEM images (figure 2.3). A Gaussian distribution curve (figure 2.4) was constructed using these results to demonstrate the

wide range of sphere lengths. The sample gave a mean of 7.0 μm and a standard deviation of 3.9 μm.
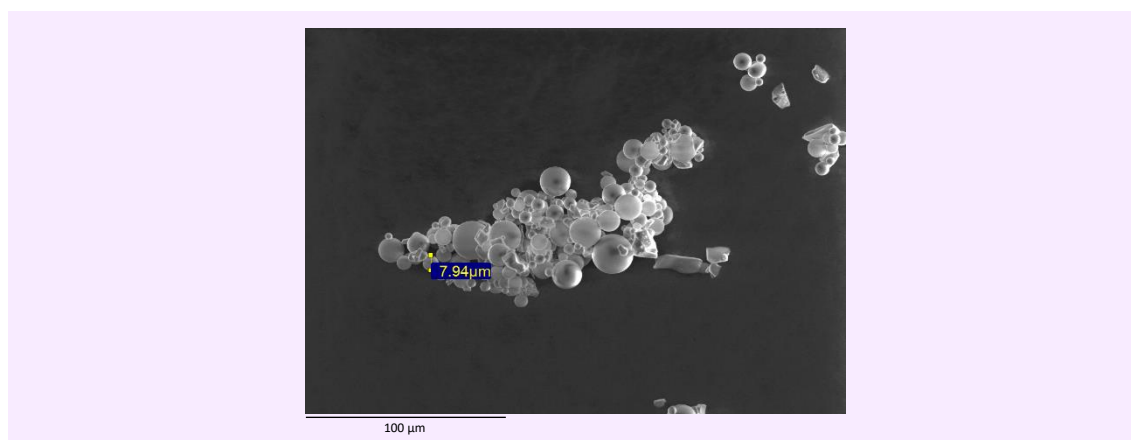


Figure 2.3. Microscopy images captured from SEM of the micro glass spheres.
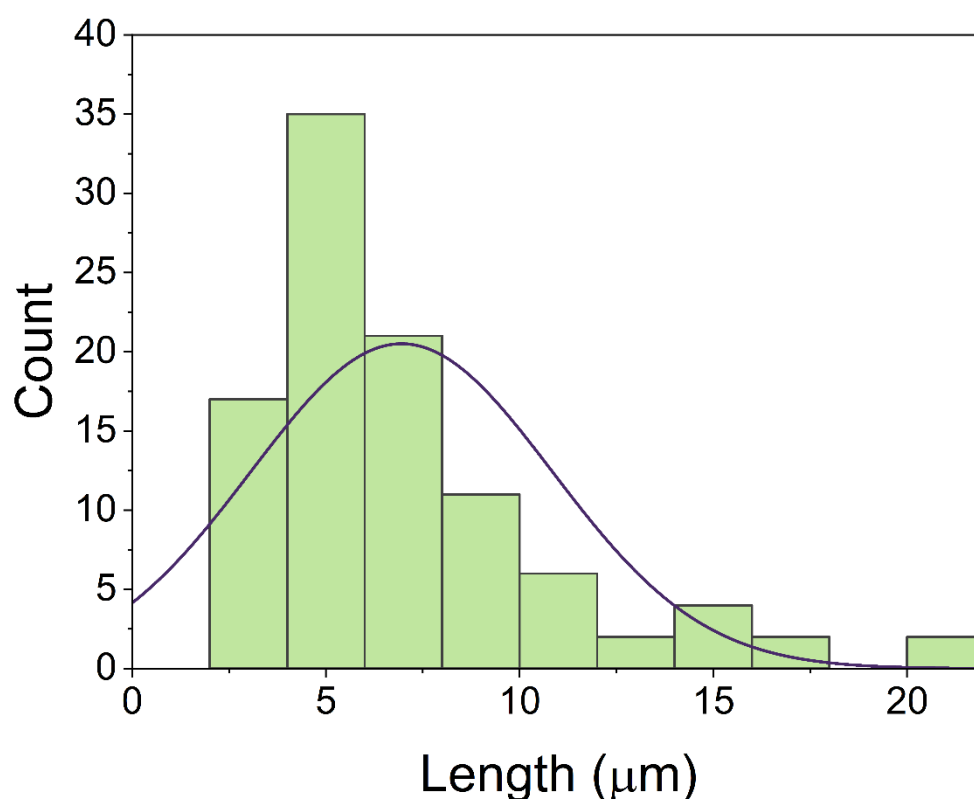


Figure 2.4. Gaussian distribution curve of the diameters of the micro glass spheres.

The micro glass spheres were put into 4 different solvents (water, TAMg buffer, tris-borate buffer and phosphate buffered saline) to check their compatibility and ability to make a stable suspension. The spheres sat on the surface of all four of the solvents. Having a solid support

that behaves as a stable suspension is important for the aptamer synthesis, these results show that these spheres are not suitable.

### 2.4.1.2. Amine Modification of the Micro Glass Spheres
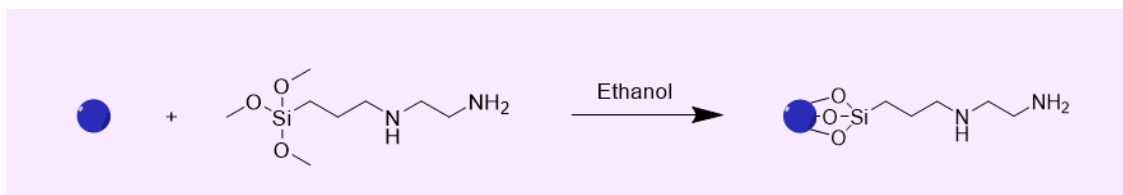


Figure 2.5. Amine modification of micro glass spheres.

The micro glass spheres were reacted with [3-(2-aminoethylamino)propyl]trimethoxysilane] to give them amine functionality, scheme shown in figure 2.5. This was done to check the spheres were able to be chemically modified. This product was separated into two sizes: ~2-10 μm and ~11-22 μm by using filter paper with pore size of 10 μm. The Kaiser test[349] was run on both samples to check for the presence of a primary amine. The Kaiser test gives a dark purple colour when positive, which was visible in the products of these reactions. Microscopy images were captured of these products (figure 2.6). Filtering the smaller sized spheres out of the product did not appear to make them significantly more uniform in size, and it did not stop their being other non-spherical items in the samples.



~11-22 μm

~2-10 μm
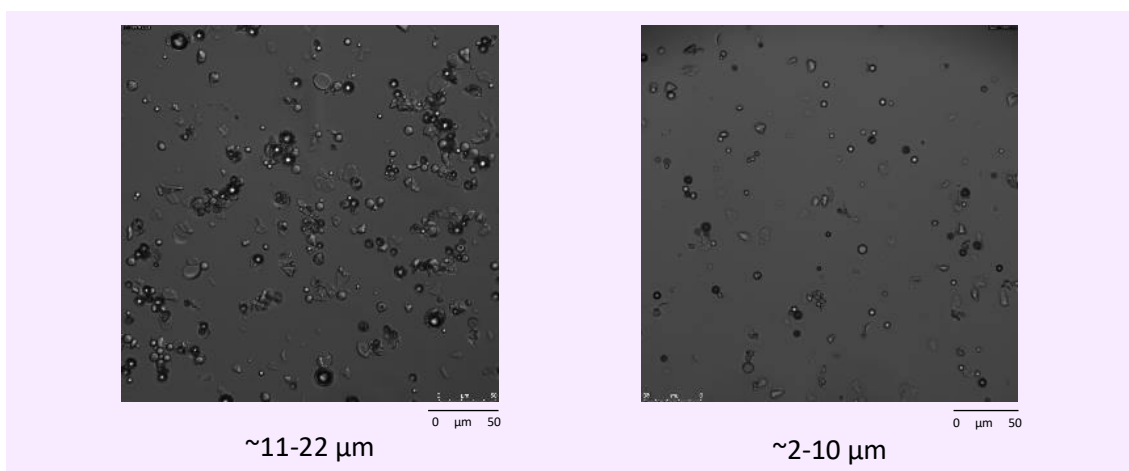
Figure 2.6.  Microscopy images captured of the amine modified micro glass spheres.

## 2.4.1.3. Fluorescent Amine Modification of the Micro Glass Spheres
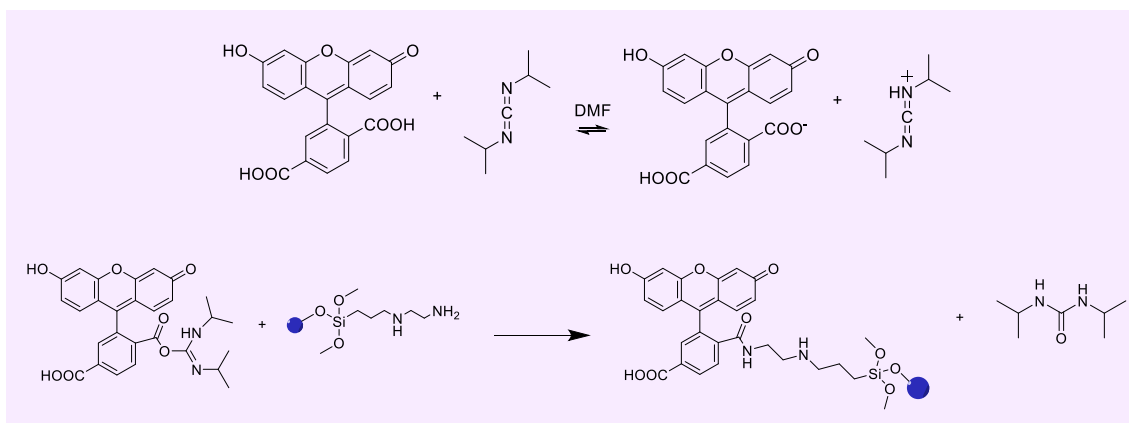


Figure 2.7.  Scheme of the 6-carboxyfluorescein modified amine micro glass spheres.

The amine modified micro glass spheres were tagged with 6-carboxyfluorescein, which is a fluorescent dye with the absorption wavelength of 495 nm and an emission wavelength of 517 nm (figure 2.7). Fluorescent microscope analysis of the tagged micro glass spheres shows the reaction was successful (figure 2.8).

Figure 2.8 shows that the fluorophore binding was only to the outside of the spheres. The background is not jet black as expected meaning the fluorescent tag is not bound tightly. Again, there are other objects in this microscope image, impurities that are not being fluorescently tagged so are potentially not silica.
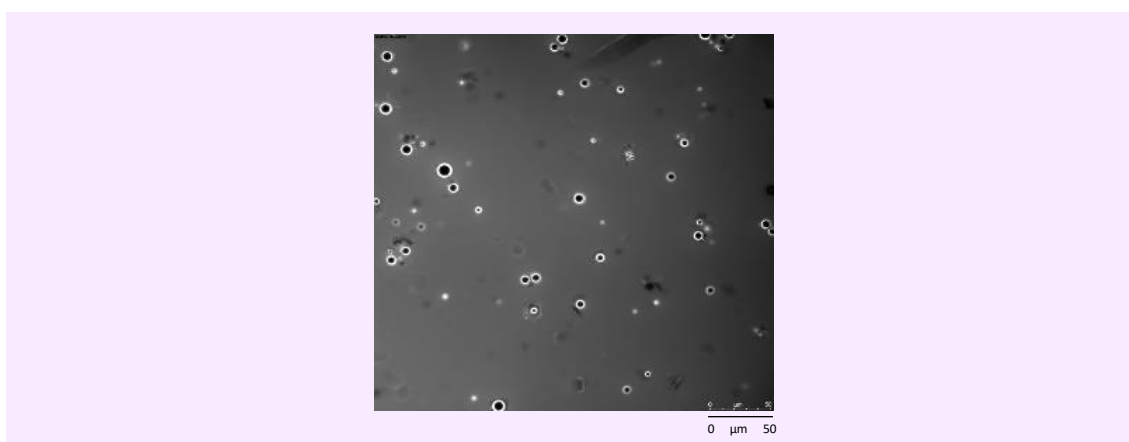


Figure 2.8.  Fluorescent Microscopy images captured of the 6-carboxyfluorescein tagged amine modified micro glass spheres.

2.4.1.4. The binding capacity of the micro glass spheres

The binding capacity of the micro glass spheres was calculated by working out how many [3-(2-aminoethylamino)propyl]trimethoxysilane] molecules could potentially bind to the surface of the averaged size micro glass sphere (11 µm) (table 2.4).

Table 2.4. Predicted Binding Capacity of the Micro Glass Spheres.

| Per       Micro Glass Spheres | Number of potential binding sites | Number of beads per gram | Mol | Potential Binding Capacity |
|---|---|---|---|---|
| Predicted | 95033177.8 | 5.40E+8 | 1.58E-16 mol | 8.5E-5 mmol/g |

To test these predictions a quantitative amine test was carried out by creating a set of free amine solution standards, running them on the Nanodrop UV-vis spectrophotometer and then plotting a calibration graph. Ninhydrin dye was used to show the concentration of the free amine in the standards by visible colour and by absorbance detection (figure 2.9).
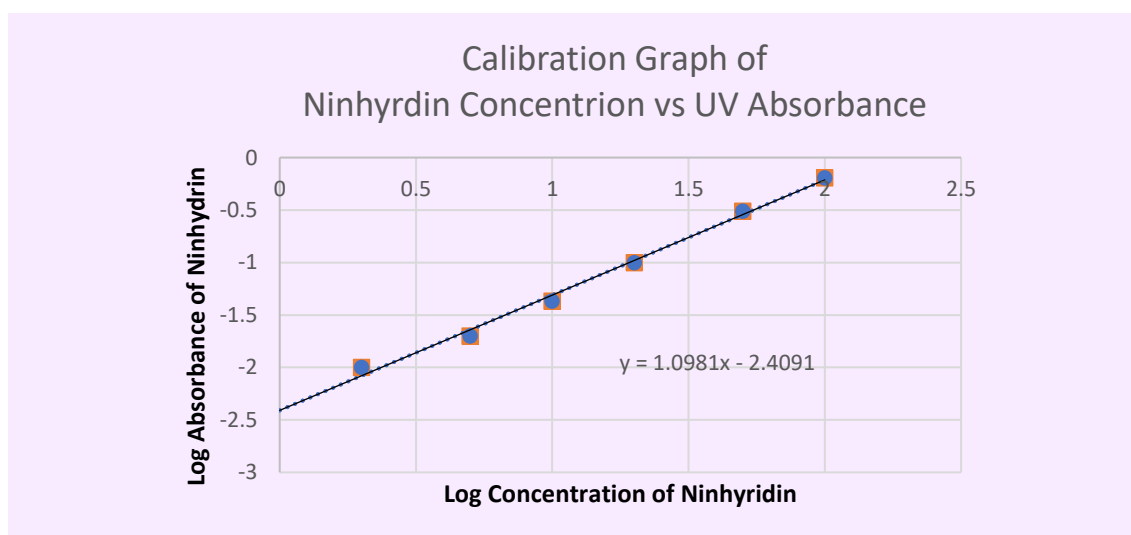


Figure 2.9. Log-Log Calibration Graph of the Quantitative Amine Test. Absorbance vs Concentration.

Using the equation of the calibration graph *y = 1.0981x - 2.4091, y* representing ninhydrin absorbance and *x* representing the concentration of free amine.

110

$$Log(0.02625) = 1.0981(logx) - 2.4091. \ x = 5.68$$

The binding capacity of the micro glass spheres is 5.7% of what was predicited (table 2.4). Table compares both results.

Table 2.5.  Predicted and Experimental Binding Capacity of the Micro Glass Spheres.

| *Per* Micro Glass Spheres | Number of potential binding sights | Number of beads per gram | Mol | Potential Binding Capacity |
|---|---|---|---|---|
| Predicted | $9.50 \times 10^7$ | $5.40 \times 10^8$ | $1.58 \times 10^{-16}$ mol | $8.5 \times 10^{-5}$ mmol/g |
| Experimental | $5.89 \times 10^6$ | $5.40 \times 10^8$ | $1.58 \times 10^{-16}$ mol | $5.29 \times 10^{-6}$ mmol/g |

In conclusion, the glass micro spheres are not compatible for this project. The experimental binding capacity is too low, especially for the limit of detection in the mass spectrometer. The larger size range could potentially be problematic when running them through the FACS as the FACS tubing is not built to allow anything over 20 µm to pass through. The aptamer selection process would also achieve more accurate results if the aptamer library has been synthesised on homogenous beads because it will give them the same chance to bind with high affinity to EGFR, as each new aptamer in theory will have the same number of strands synthesised on their solid support. Having a bead size range of 2.4-21.5 µm may disadvantage new aptamers that are synthesised on the smaller beads in this range. A different solid support will be tested and characterised for all future work.

## 2.4.2. TentaGel® M NH₂ Monosized Amino TentaGel Microspheres

### 2.4.2.1. Characterisation of TentaGel® M NH₂ Monosized Amino TentaGel Microspheres
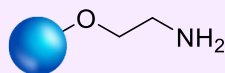


Figure 2.10.  Structure of TentaGel® M NH₂ Monosized Amino TentaGel Microspheres

TentaGel® M NH₂ Monosized Amino TentaGel Microspheres have a polystyrene backbone with a polyethylene glycol (PEG) spacer attached via alkyl linkage[351]. The structure of the surface of the TentaGel microspheres is shown in figure 2.10. This linkage is not sensitive to acids or bases, so will be suitable to be used in phosphoramidite chemistry. The TentaGel microspheres are mono-sized, mono-spherical and are 10 µm when dry.

A Kaiser test was run on the TentaGel microspheres, resulting in a positive result as shown by the production of a purple colour, indicating the presence of a free primary amine. The TentaGel microspheres were observed using the optical microscope to compare the size distribution with the previous micro glass spheres. Figure 2.11 shows that they are homogenous in size and that when the microspheres are swollen in water their diameter expands a little above 10 µm.
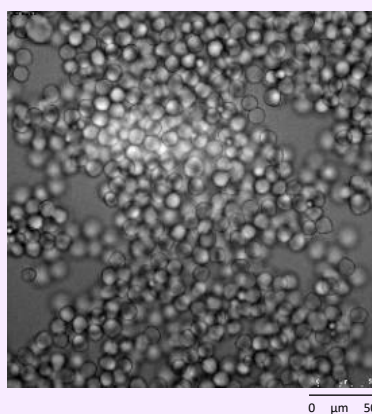


Figure 2.11.  Microscopy images captured of the TentaGel® M NH₂ Monosized Amino TentaGel Microspheres.

A sample of the TentaGel microspheres diameters were measured using SEM images (figure 2.11). A Gaussian distribution curve (figure 2.12) was made from these results to demonstrate

the range of sphere lengths. The sample gave a mean of 12.1 μm and a standard deviation of 0.66 μm.
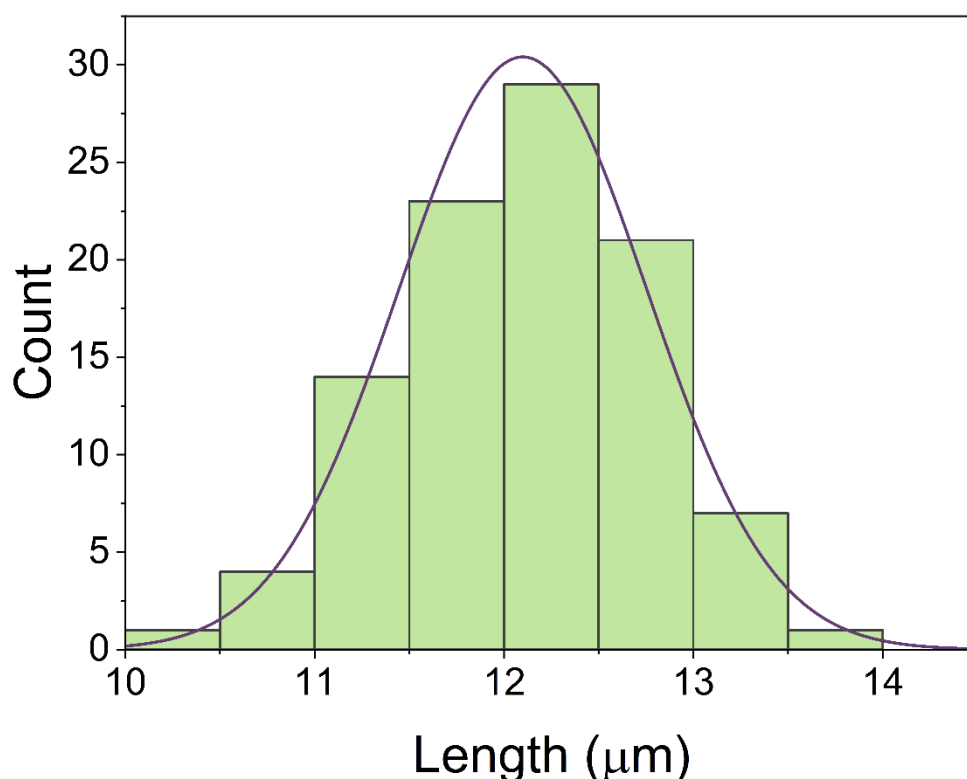


Figure 2.12. Gaussian distribution curve of the diameters of the TentaGel microspheres.

The size distribution of the TentaGel microspheres is much smaller than the micro glass spheres which makes them a better solid support for the synthesis of the aptamer library.

2.4.2.2. Binding Capacity of TentaGel® M NH$_2$ Monosized Amino TentaGel Microspheres

The binding capacity of the TentaGel microspheres is 0.28 mmol/g, as experimentally determined by Rapp Polymere. When these results were compared with the modified micro glass spheres, the TentaGel microspheres have a much better binding capability by a factor of 10$^5$ (table 2.6).

Table 2.6. Binding Capacity of the Micro Glass Spheres Vs TentaGel® M NH₂ Monosized Amino TentaGel Microspheres.

| | Number of potential binding sites per sphere | Number of beads per gram | Mol | Potential Binding Capacity |
|---|---|---|---|---|
| TentaGel Microspheres | $1.17 \times 10^{11}$ | $1.49 \times 10^{9}$ | $1.94 \times 10^{-13}$ mol | $2.80 \times 10^{-1}$ mmol/g |
| Micro Glass Spheres | $5.89 \times 10^{6}$ | $5.40 \times 10^{8}$ | $1.58 \times 10^{-16}$ mol | $5.29 \times 10^{-6}$ mmol/g |

After running the same characterisation tests on both solid supports and comparing the binding capacities, the TentaGel microspheres are found to be the better solid support and will be used in the next stage of the project. Their homogenous nature and their higher binding capacity makes them the better candidate along with the elasticity of them making them more compatible for tubing in the FACS.

### 2.4.2.3 Synthesising Oligonucleotides on to TentaGel® M NH₂ Monosized Amino TentaGel Microspheres

MinE07ARP:5'-rGrGrAfCrGrGrAfUfUfUrArAfUfCrGfCfCrGfUrArGrArArArArGfCrAfUrGfUfCrArAr ArGfCfCrGrGrArAfCfCrGfUfCfC-3'.

MinE07 was synthesised on the TentaGel Microspheres to see if this solid support was suitable for oligonucleotide synthesis. All the aptamer synthesis will follow the same steps and synthetic reaction scheme (figure 2.13). Step 1: Deblocking. 3'-5' nucleotide addition first requires removal of the dimethoxytrityl (DMT) protecting group on the 5' carbon of the first nucleotide. This is accomplished by introducing a deblocking reagent, which leaves the first nucleotide with a free 5'-OH group that can then react with the next nucleotide. The trityl monitor detects the

acid-labile DMT protecting group removal. The amount of DMT released from the column is an indication of the success of the addition of the last monomer to the growing oligonucleotide chain. The DMT group is orange in the acidic deblocking solution which the trityl monitor is set to detect as it passes through a flow cell. There is a light source on the trityl monitor emitting at 470 nm which DMT absorbs. Data from the trityl monitor is displayed in a bar graph. Each bar is an indication of the success of the previous coupling. It represents the cumulative photometric data collected from the trityl monitor during the deblocking step. Response value for the highest trityl bar (excluding the first bar) that is a qualitative indication of coupling efficiency. The overall appearance of the bar graph can appear rough. This irregularity can be caused by deblock solution out-gassing within the flow cell or the DMT deblocking rate differences for various monomers.

Step 2: Coupling. As a new nucleotide is introduced to the first nucleotide, it reacts with a weak acid to form a phosphoramidite intermediate. This allows the nucleotide to interact with the unblocked -OH group on the 5' end of the receiving nucleotide, producing a covalent attachment by a phosphite triester bond.

Step 3: Capping. There is never 100% coupling efficiency. Some nucleotides will not couple with the added nucleotide, leaving a free 5' -OH group. If not capped, this group can couple to the nucleotide added in the next cycle, which will produce an oligonucleotide with a deletion in its sequence. To avoid this happening, a capping step is used to prevent any uncoupled molecules from further reaction.

Step 4: Oxidation. The newly bound nucleotide must be stabilised. This is done by adding a mixture of iodine and water. The mixtures oxidise the phosphite into phosphate, resulting in a stable phosphotriester bond between the oligonucleotide and its new nucleotide.

Step 5: Phosphoramidite addition repeats until the desired sequence is complete. Once the last nucleotide is attached, oligonucleotides undergo a removal of protecting groups that are still attached. After these groups are removed, oligonucleotides are cleaved from the solid support and desalted to remove any additional contaminants.
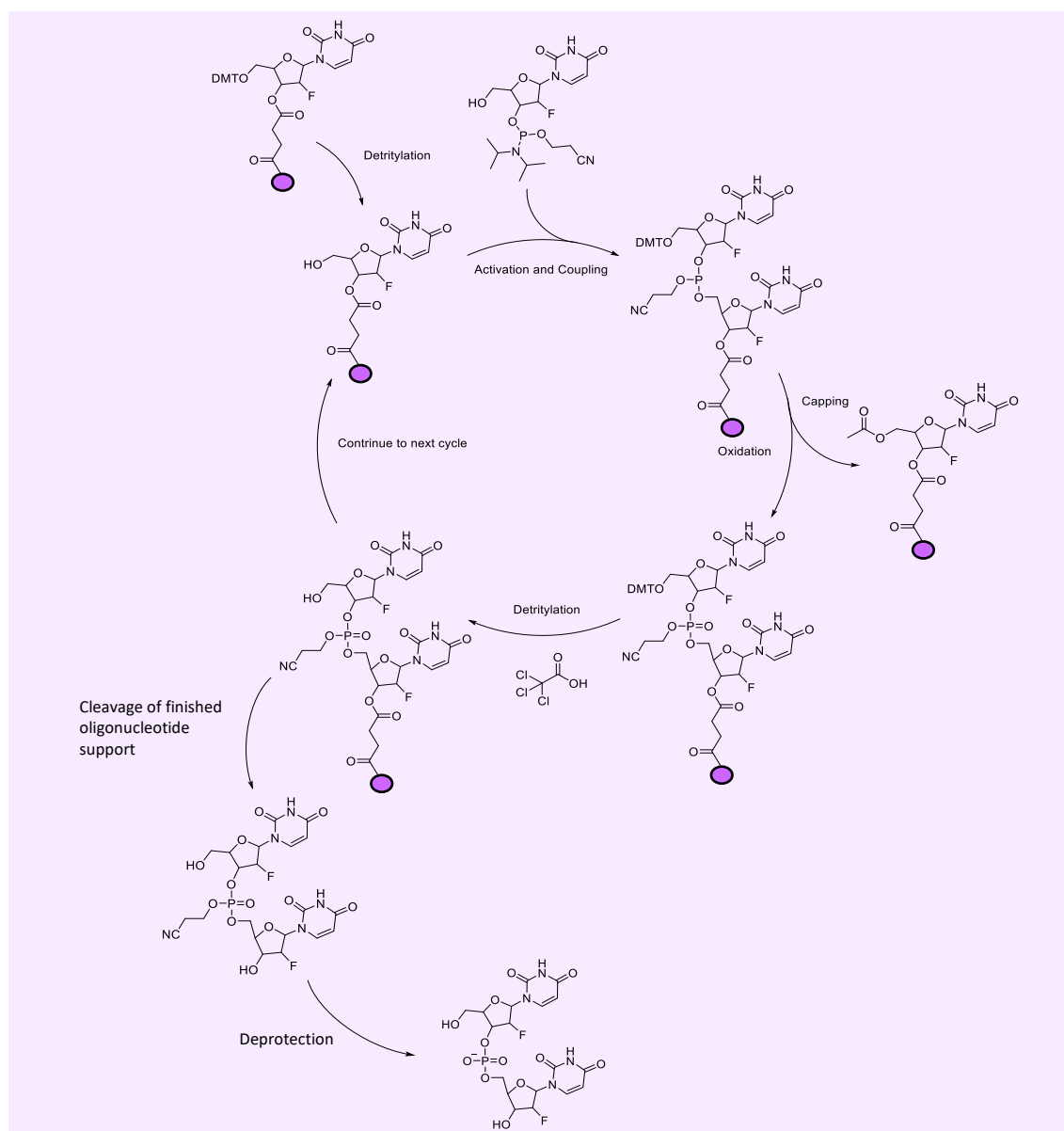


Figure 2.13. Aptamer Synthesis Synthetic Reaction Scheme.

The Expedite™ DNA synthesiser was used to generate the MinE07 parent aptamer (MinE07ARP) on the TentaGel microspheres (0.2 µmol). The first coupling was not successful, which is believed to be because of the $-NH_2$ group on the TentaGel microspheres was not coupling well

with the first base. A linker was chosen to change the functional group to an –OH which will

couple more efficiently.
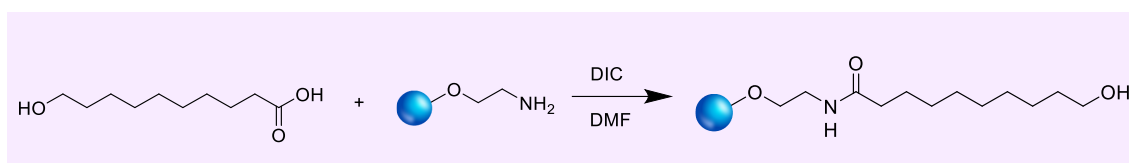


Figure 2.14. Addition of 10-hydroxydecanoic linker to TentaGel® M NH₂ Monosized Amino TentaGel Microspheres.

10-hydroxydecanoic acid was the chosen linker to attach to the TentaGel microspheres (figure

2.14). Ninhydrin dye was used to verify that the reaction was successful. The results of the Kaiser

test were negative (yellow in colour), so the reaction was successful as no primary amines were

shown to be present. MinE07ARP was synthesised onto the –OH modified TentaGel

microspheres. The protocol for coupling was extended 10x more than a standard DNA 1 µmol

protocol. This was to allow more time and reagents to permeate through the polystyrene in the

TentaGel microspheres. The protocol change can be found in figure 2.15, highlighted below.

```
Cycle A (dAdenosine) 1umol x10 time ARP
/* ------------------------------------------------------------------------ */
/*        Function              Mode  Amount Time(sec)        Description    */
/*                                    /Arg1  /Arg2                           */
/* ------------------------------------------------------------------------ */
$Deblocking
 144 /*Index Fract. Coll. */ NA      10      0   "Event out ON"
   0 /*Default            */ WAIT     0     15   "Wait"
 141 /*Trityl Mon. On/Off */ NA      10     10   "START data collection"
  16 /*Dblk               */ PULSE  300      0   "Dblk to column"
  16 /*Dblk               */ PULSE  500    490   "Deblock"
  16 /*Dblk               */ PULSE  500    390   "Deblock"
  38 /*Diverted Wsh A      */ PULSE  400      0   "Flush system with Wsh A"
 141 /*Trityl Mon. On/Off */ NA       0     10   "STOP data collection"
  38 /*Diverted Wsh A      */ PULSE  400      0   "Flush system with Wsh A"
 144 /*Index Fract. Coll. */ NA      20      0   "Event out OFF"
$Coupling
   1 /*Wsh                */ PULSE   50      0   "Flush system with Wsh"
   2 /*Act                */ PULSE   50      0   "Flush system with Act"
  18 /*A + Act            */ PULSE   50      0   "Monomer + Act to column"
  18 /*A + Act            */ PULSE   20    160   "Couple monomer"
   2 /*Act                */ PULSE   30    240   "Couple monomer"
   1 /*Wsh                */ PULSE   70    560   "Couple monomer"
   1 /*Wsh                */ PULSE   80      0   "Flush system with Wsh"
$Capping
  12 /*Wsh A              */ PULSE  200      0   "Flush system with Wsh A"
  13 /*Caps               */ PULSE   80      0   "Caps to column"
  12 /*Wsh A              */ PULSE   60    150   "Cap"
  12 /*Wsh A              */ PULSE  140      0   "Flush system with Wsh A"
$Oxidizing
  15 /*Ox                 */ PULSE  150      0   "Ox to column"
  12 /*Wsh A              */ PULSE  150      0   "Flush system with Wsh A"
$Capping
  13 /*Caps               */ PULSE   70      0   "Caps to column"
  12 /*Wsh A              */ PULSE  300      0   "End of cycle wash"
```

Figure 2.15. 1 µmol 10x Protocol for the synthesis of MinE07ARP on TentaGel Microspheres.

This method was still unsuccessful, so the solvent the bases were dissolved in was changed from acetonitrile to DCM (better permeability of the TentaGel), this worked a lot better and the full length MinE07ARP was synthesised. The trityl monitor progress report is shown (figure 2.16).
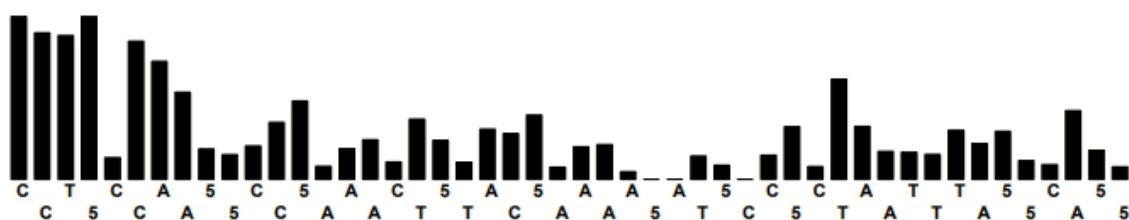


Figure 2.16. Trityl Monitor Report of the synthesis of MinE07ARP on modified TentaGel® M NH₂ Monosized Amino TentaGel Microspheres.

The step wise coupling efficiency is 96.9% as shown in the synthesis report (figure 2.17), To get the overall average yield: $0.969^{48} = 0.221$. This gives a 22.1% overall synthetic yield, which is good enough to carry forward for more experiments.

Figure 2.17. Aptamer Synthesis Report of the Synthesis of MinE07ARP on modified TentaGel® M NH₂ Monosized Amino TentaGel Microspheres. Monomer T as uridine, monomer 5 as guanine.

## 2.4.3. Flow Cytometry of fluorescently tagged TentaGel® M NH₂ Monosized Amino TentaGel Microspheres

2.4.3.1. Synthesising Fluorescently Tagged TentaGel® M NH₂ Monosized Amino TentaGel Microspheres

To check the compatibility of the TentaGel microspheres in the FACS and to analysis the sensitivity of its fluorescent detector a range of fluorescently tagged TentaGel microspheres were synthesised. 6-carboxyfluorescein was the first fluorescent tag used, TentaGel microspheres with different percentages of 6-carboxyfluorescein (100%, 73% and 36%) attached were synthesised (figure 2.18).
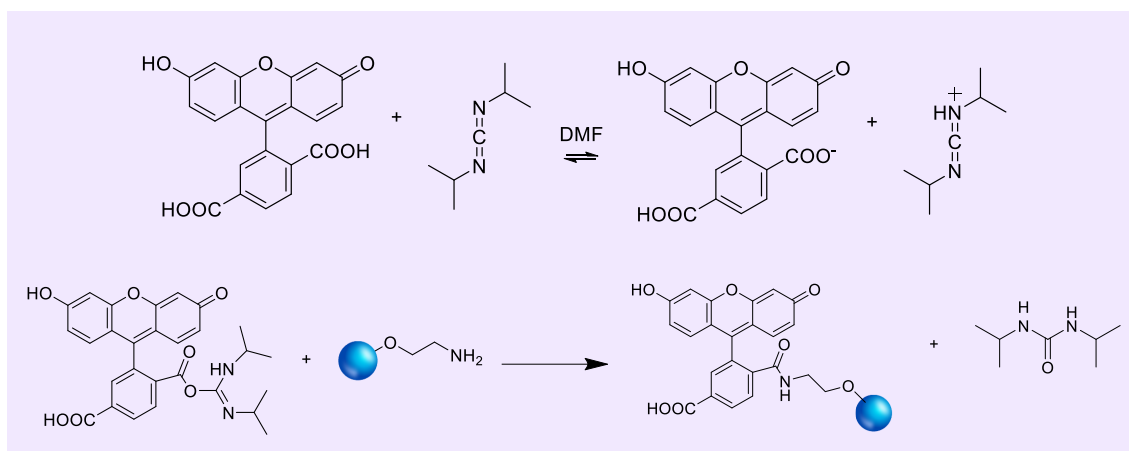


Figure 2.18.  Scheme of the 6-carboxyfluorescein taggaed TentaGel® M NH₂ Monosized Amino TentaGel Microspheres.

119

The 100% 6-carboxyfluorescein tagged TentaGel microspheres were analysed by optical microscope, images (figure 2.19). The images show that the binding of the 6-carboxyfluorescein was successful as the fluorescent is very bright and the background is showing no fluorescent activity meaning the 6-carboxyfluorescein tag is attached to the microspheres and not in the supernatant.
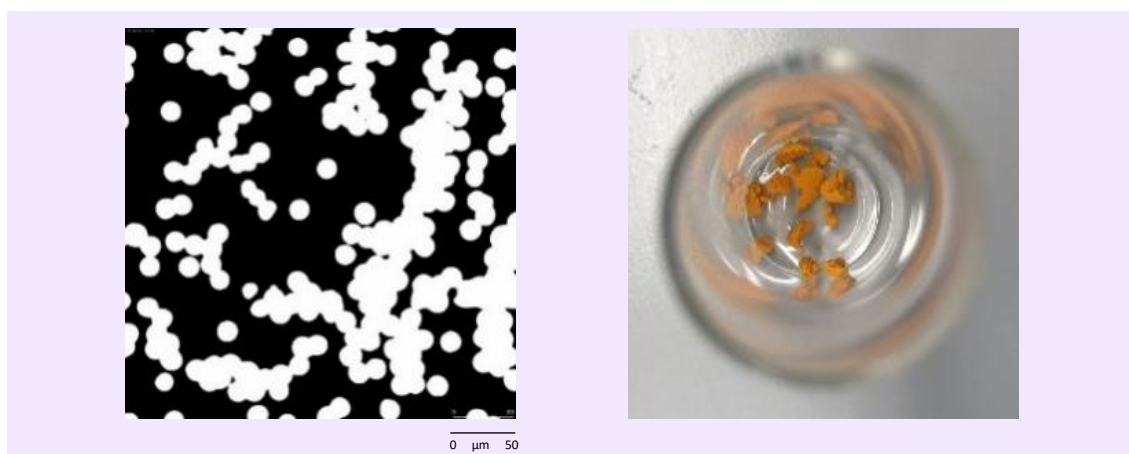


Figure 2.19. Microscopy images captured of the 6-carboxyfluorescein tagged TentaGel® M NH$_2$ Monosized Amino TentaGel Microspheres.

Rhodamine B was also tried as this uses a different laser on the FACS. TentaGel microspheres with 100%, 73% 36%, 17% and 1.7% tag attached were synthesised. The 100% rhodamine B tagged TentaGel microspheres were analysed by optical microscope, images (figure 2.21).
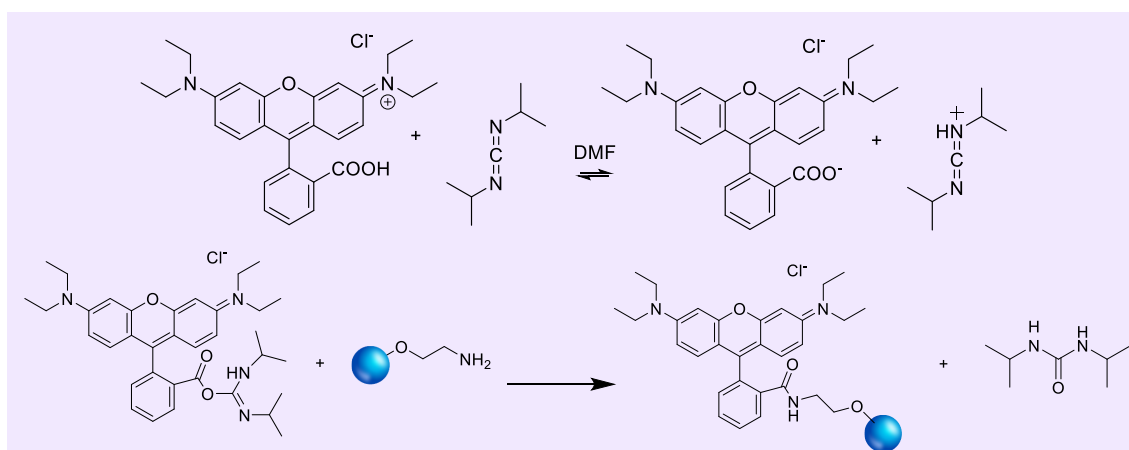


Figure 2.20. Scheme of the Rhodamine B tagged TentaGel® M NH$_2$ Monosized Amino TentaGel Microspheres.

The images show that the binding with the rhodamine B was also successful as the fluorescent is very bright and the background is showing no fluorescent activity meaning the rhodamine B tag is attached to the microspheres and not in the supernatant.



Figure 2.20. Microscopy images captured of the Rhodamine B tagged TentaGel® M NH$_2$ Monosized Amino TentaGel Microspheres.

2.4.3.2. Flow Cytometry Analysis of Fluorescently Tagged TentaGel® M NH$_2$ Monosized Amino TentaGel Microspheres

The fluorescent microsphere samples TGRhodB100, TGRhodB73, TGRhodB36, TGRhodB17, TGRhodB1.7, TGCFluor100, TGCFluor73 and TGCFluor36 were put through the FACS to detect the level of fluorescence compared with the plain microspheres. The aim was to see if the difference in the level of fluorescence on the different microspheres was visibly distinguishable. The 6-carboxyfluorescein tagged microspheres were detected in the FACS using the laser with the excitation of 488 nm and emission of 517/17 nm. The FACS data (figure 2.22) shows there is a difference in fluorescence between the tagged microspheres and the non-tagged microspheres. The histogram shows that for TGCFluor100, TGCFluor73 and TGCFluor36 there was not a huge visual different between the three tagged samples. Potentially this is because the reactive is not working as well as would have wanted, even though the microscopy images show a successful conjugation (figure 2.19).
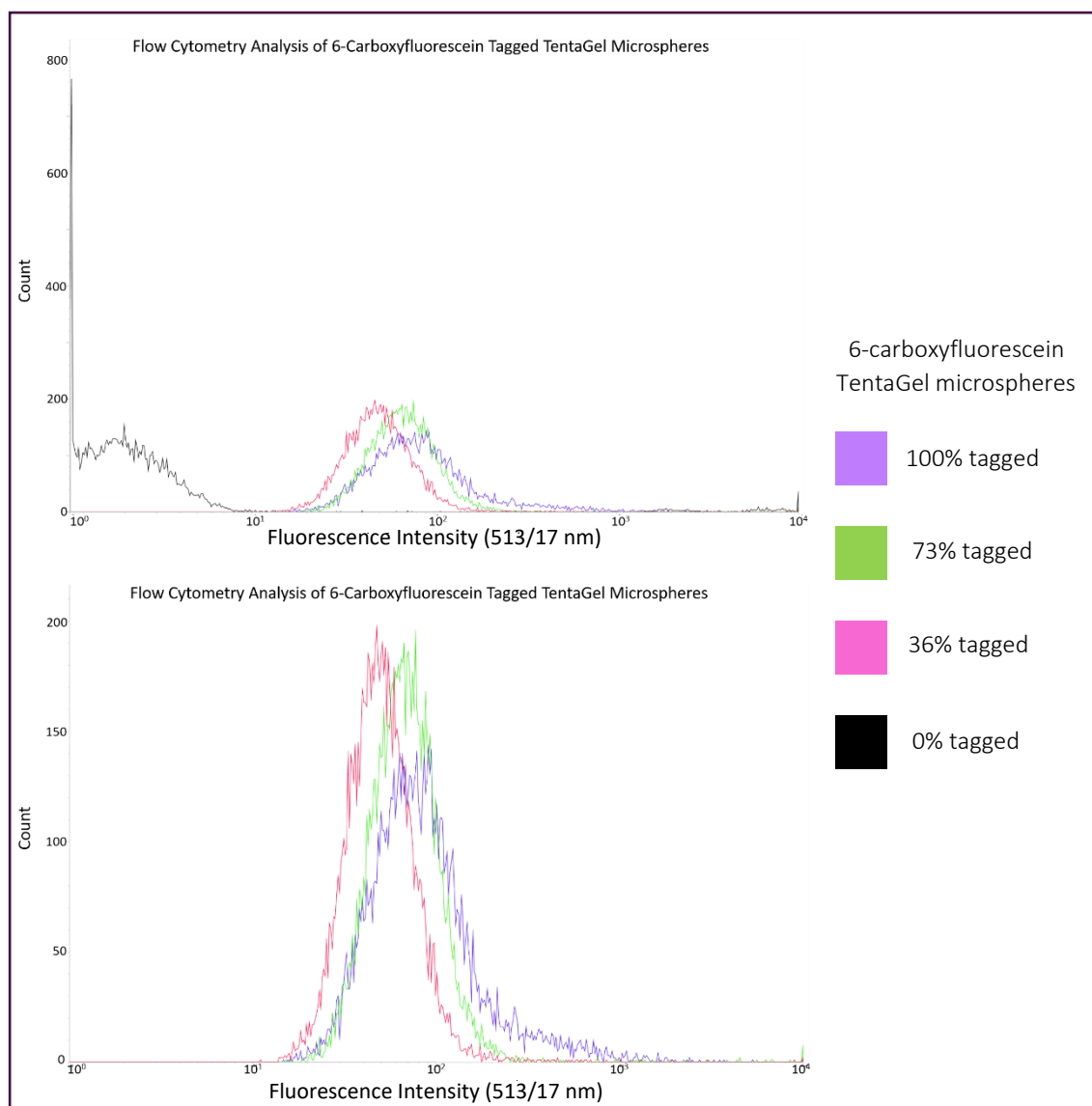
Figure 2.22.  FACS data of 6-carboxyfluorescein tagged TentaGel® M NH$_2$ Monosized Amino TentaGel Microspheres.

The rhodamine B tagged microspheres were detected in the FACS using the laser with the excitation of 561 nm and the emission of 585/29 nm. Figure 2.23 showing the FACS data shows that there is a difference in fluorescence between the 5 tagged microsphere samples and the non-tagged microspheres. This data proves that the FACS can distinguish between levels of fluorescence using rhodamine B.
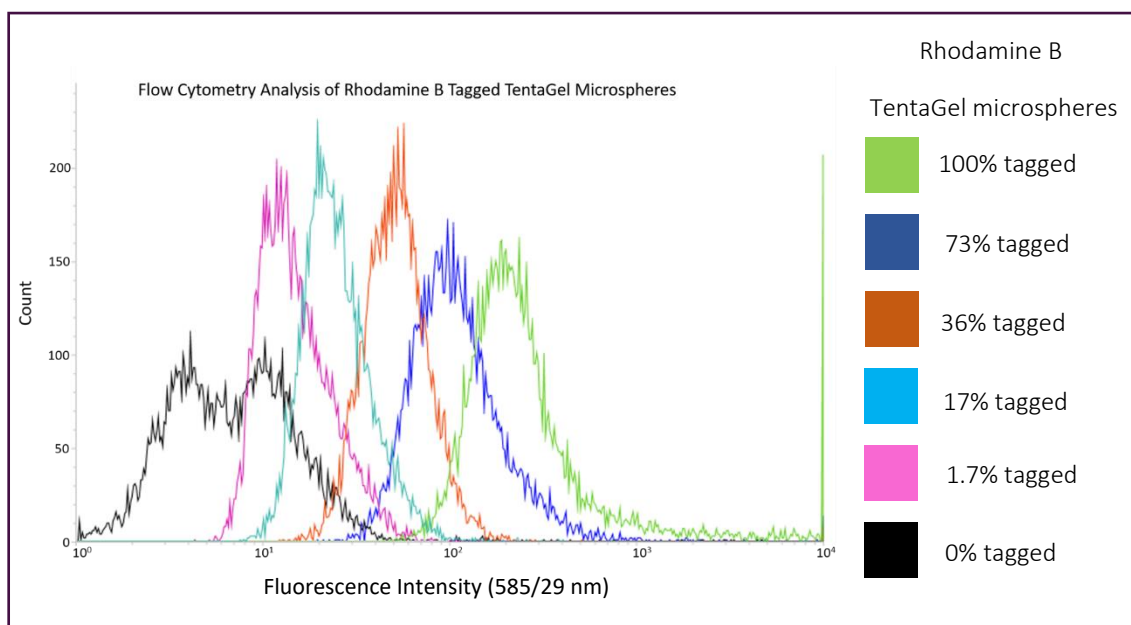
Figure 2.23. FACS data of Rhodamine B tagged TentaGel® M NH$_2$ Monosized Amino TentaGel Microspheres.

Doublet at 0% tagged caused by two microspheres going through the laser simultaneously.

### 2.4.4.3. Flow Cytometry Analysis of MinE07ARP Tagged TentaGel® M NH$_2$ Monosized Amino TentaGel Microspheres.

The MinE07ARP tagged TentaGel Microspheres were put through the FACS to see if they displayed any fluorescence. They were compared with plain microspheres, TGRhodB100 and TGCFluor100 microspheres. This was done to check that there was not any fluorescence intensity higher than the plain microspheres. They were detected in the FACS using the laser with the excitation of 488 nm and the emission of 517/17 nm (figure 2.24). The MinE07ARP microsphere signal overlapped with the plain microsphere signal demonstrating that they did not fluoresce.
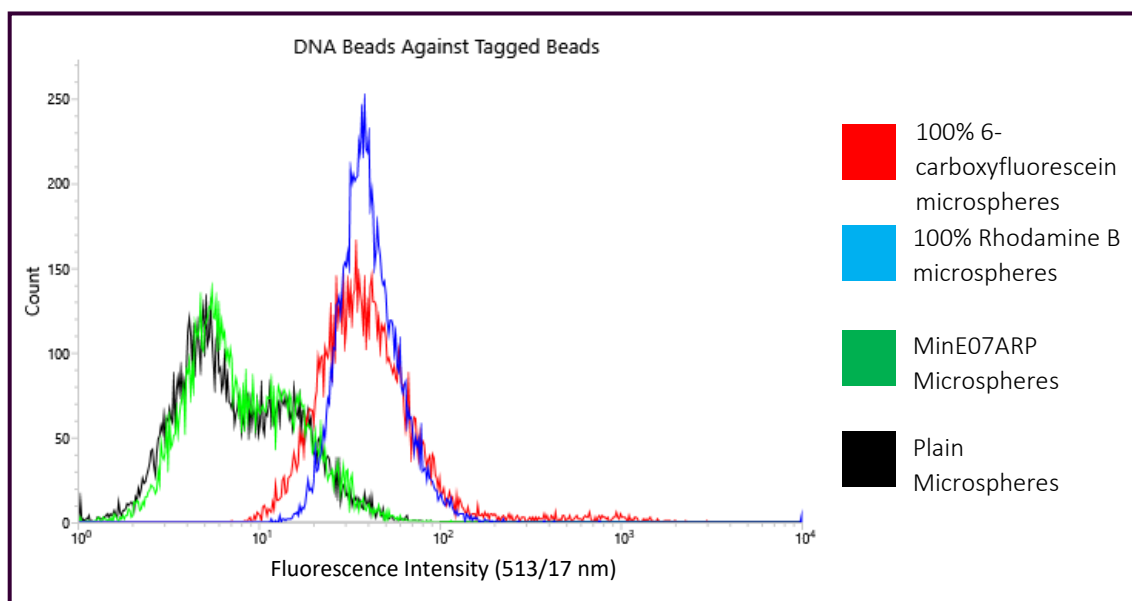
Figure 2.24. FACS data of MinE07ARP tagged TentaGel® M NH₂ Monosized Amino TentaGel Microspheres. Doublet caused by two microspheres going through the laser simultaneously.

## 2.5. Conclusion

In conclusion, the TentaGel Microspheres have been experimentally validated as a suitable solid support for the synthesis of the MinE07 aptamer library followed by flow cytometry. The TentaGel microspheres can have fluorescent tags bound to them, visible on the optical microscope. Different amounts of fluorescence tagged on to the TentaGel microspheres is distinguishable in the FACS using the rhodamine B tag, which proves that sorting by fluorescence intensity using flow cytometry methods will work with these microspheres on the FACS. The carboxyfluorescein microspheres should be re-synthesised and tried again on the FACS. The microspheres are capable of being chemically modified to include a linker allowing for the MinE07 aptamer to be successfully synthesised on the surface of the microspheres.

Finding a successful solid support is vital to this method, having one that is chemically compatible for aptamer synthesis and works in the FACS is essential.

# Chapter 3

## Nucleic Acid Modification Design and Synthesis

# Chapter 3

## 3.1. Introduction

### 3.1.1. Nucleic Acid Modification

#### 3.1.1.1. Nucleic Acid Chemistry

Nucleic acids are fundamental molecules of life, carrying the genetic information of all living organisms. Their main function is to store and transmit genetic information with sufficient stability, flexibility and accuracy to the next generation of each living organism .[352] The chemistry behind these structures was first uncovered by Franklin, Wilkins, Watson and Crick in 1953 who brought to light the nature of the double helical structure of deoxyribonucleic acid (DNA).[353] Nucleic acids take part in predictable and programmable interactions, they can easily denature and then refold themselves and they are water-soluble. These characteristics allow them to be used for a number of applications including therapeutics, diagnostics, biotechnology, nanotechnology, synthetic biology and material science.[354]
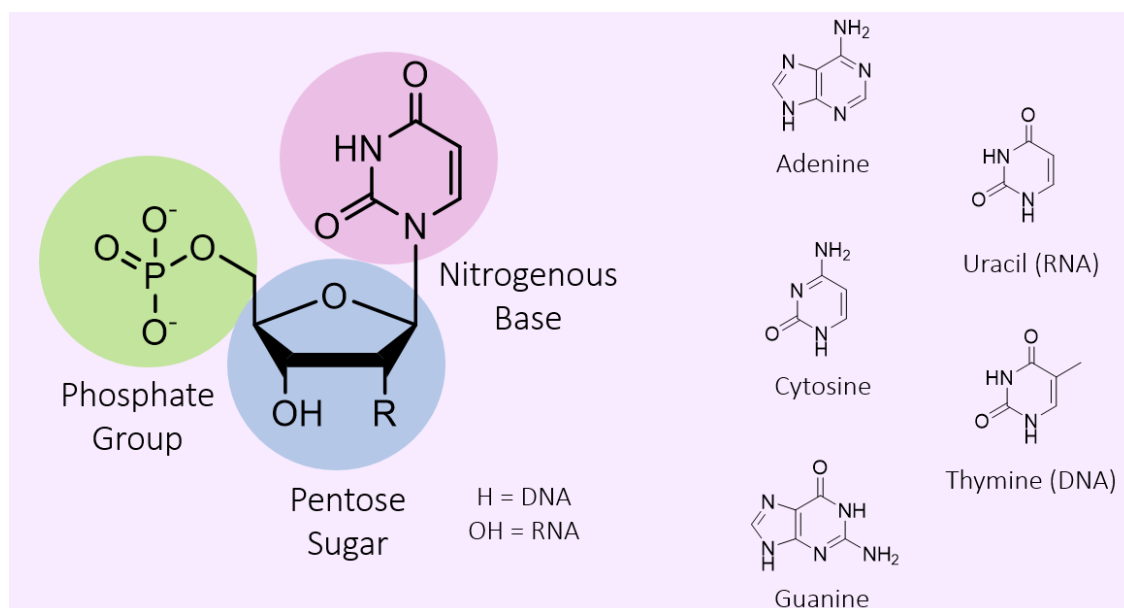


Figure 3.1. Nucleotide Structure.

A nucleotide consists of a phosphate, a pentose sugar and a nitrogenous base.[355] In nature, the nitrogenous bases are made up of purines and pyrimidines. Purine nucleotides consist of adenine and guanine. Pyrimidine nucleotides consist of cytidine, uridine (only for RNA) and thymidine (only for DNA),[356] the structures of these can be found in figure 3.1. The difference between RNA and DNA is the atoms at the C-2 position. In DNA there is just a hydrogen (H) atom and in RNA there is a hydroxyl (OH) group instead. The phosphate group attaches by a phosphoester bond making the nucleoside turn into a nucleotide. The phosphoester bond is linked by 5'-hydroxyl group of the pentose sugar and a phosphate group. Each of these three components displayed in figure 3.1 play their own individual and vital role in the structural design, synthesis and modifications of oligonucleotides.[357]

### 3.1.1.2. Nucleic Acid Modifications

Over the last 40 years there has been an increase in the use of nucleic acids as therapeutics.[358] These are normally found in the form of aptamers[359] (see chapter 1), small interfering RNAs,[360] antisense oligonucleotides (ASOs),[361] microRNAs[362] and mRNA vaccines.[363] Nucleic acids allow for a unique therapeutic intervention such as gene expression,  to generate a specific mRNA or protein that cannot be readily achieved by small molecule or biological drugs.[354] However, the natural chemistry of nucleic acids prevents them from reaching their full potential in medicine because of their limited chemical diversity and their poor biological and chemical stability.[354]

To try to reach the full potential of nucleic acids, researchers have designed various modifications at the different points of the structure. In small molecule drug discovery, it is normal practice to have a wide range of structural diversity to produce libraries of potential compounds. The creation of this type of diverse library is to make structural, functional and property refinements.[364] Structural differences define functional diversity in oligonucleotides. The types of diversity can be split into five different structural groups (figure 3.2).[365]

1. Nucleobase or sequence diversity: the change in nucleobase which then changes the whole sequence. Modification of nucleobases has a large effect on the helix formation, especially when there is substitution of their hydrogens.[366]

2. Nucleoside diversity: the difference in the chemical structure of the nucleosides which then varies the oligonucleotide.

3. Backbone diversity: the variation in the linkage structure between the nucleotides, including differences in the phosphorothioate linkages or its chiral orientation.

4. Architectural diversity: differences in design and the arrangement and order of the nucleotides in the oligonucleotides.

5. Macromolecular structure diversity: the difference in structure and topology of the complete molecule.[367]

6. The properties of the oligonucleotide are determined by the position of each structural unit in the situation of the complete molecule. The macromolecular structural diversity considering the overall structure and so can relate to the shape of the structure, the way the chain folds and its aptameric binding properties. With one single change, it will be possible to attain many different shapes and binding orientations, each representing different energy levels.[367]
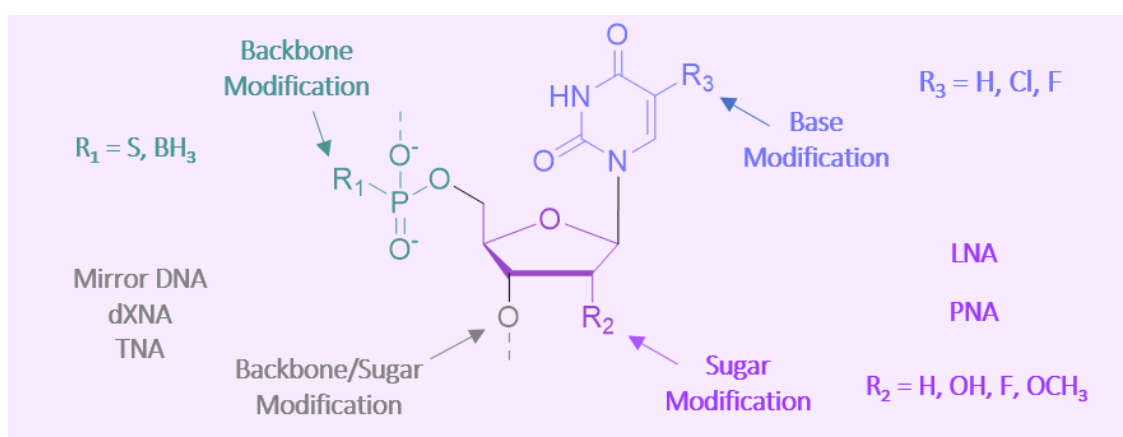


Figure 3.2. Examples of Nucleic Acid Modification Positions.

The approved anti-VEGF aptamer drug, pegaptanib, has modifications to the sugar, having either an OH, F or $OCH_3$ in the 2' position (Figure 1.12, chapter 1).[368]

### 3.1.1.3. Fluorine in Medicinal Chemistry

Fluorine atoms do not cause significant steric constraints since the fluorine atom is compact. Fluorine makes strong bonds to other elements which creates good chemical stability. Changing a hydrogen for a fluorine can cause change in the bond polarity, solubility and the possibility of creating hydrogen bonds with another molecule.[369] C-F bonds can have a bond dissociation of up to 130 kcal/mol whereas C-H bonds are somewhere around 105 kcal/mol.[370] Fluorinated compounds are useful in material chemistry and in medicinal chemistry. Many pharmaceuticals and biologically active molecules contain fluorine which has been shown to be essential for their activity.[371]

Of all drugs in pharmaceutical production, 20-25% of them contain at least 1 fluorine atom. This is fascinating considering that organofluorine compounds are generally not found in natural products.[371] One of the first fluorinated drugs was 5-fluorouracil which is an antineoplastic agent (figure 3.3). It is an antimetabolite and was first synthesised in 1957.[372] It has high anticancer activity by inhibiting the enzyme thymidylate synthase which then prevents the cellular synthesis of thymidine.
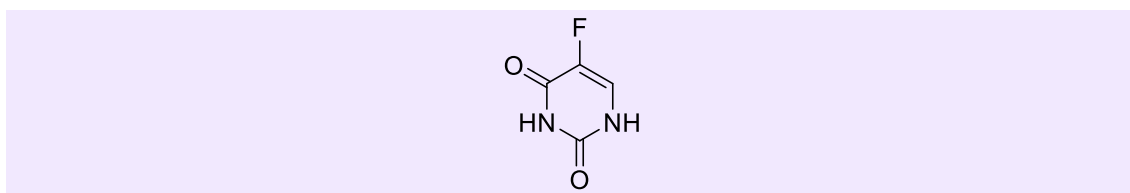


Figure 3.3. Chemical structure of 5-fluorouracil.

Fluorine substitution is often used in medicinal chemistry because it improves metabolic stability, bioavailability and protein-ligand interactions.[373] This strategy of using fluorine substitution in drug design has led to production of some key drugs on the market. Drugs such

as Fluoxetine[374] (an antidepressant), Faslodex[375] (anticancer), Flurithromycin[376] (antibacterial) and Efavirenz[377] (antiviral). These four drugs cover a wide range of different medical areas, which have all benefited from fluorine chemistry (figure 3.4).[371]
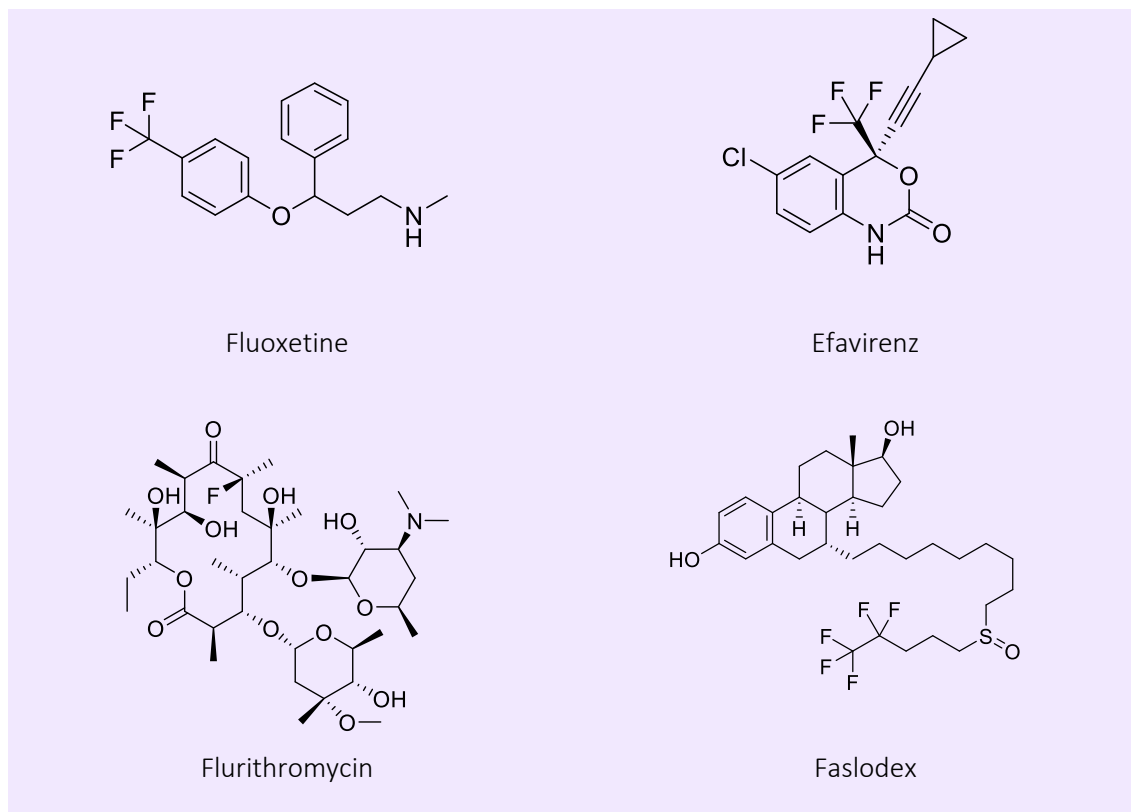


Figure 3.4. Chemical structure of fluorine containing drugs.

The uridine molecule that will be modified during this project will have a fluorine at the 2' position, 2'-fluoro-2'-deoxyuridine, as the parent MinE07 aptamer has this structure. This will be carried forward from previous research as it has been proven that this modification helps with stability and prevents degradation.

## 3.1.2. Types of Chemistry used for Nucleotide Modification

### 3.1.2.1. Suzuki-Miyaura Cross Coupling Reaction

The discovery of reactions using palladium catalysts forming carbon-carbon bonds has been important in the development of synthetic organic chemistry.[378] The advantages of the Suzuki

Miyaura cross coupling reaction are that it has high region- and stereoselectivity, it works with a large variety of functional groups, and the conditions are mild.[379] Miyaura's team worked on the coupling of alkylboron compounds with different organic halides with the catalysis $PdCl_2.CH_2Cl_2$ in the presence of a base for example $K_2CO_3$ or $K_3PO_4$ (figure 3.5).[380] These reactions have proved extremely useful as a method for synthesis of biologically active and natural compounds.[381] Use of boronic acids and boronate esters in these reactions also have many limitations of use such as high expense, purification difficulties and sometimes air and moisture sensitivity.[382]
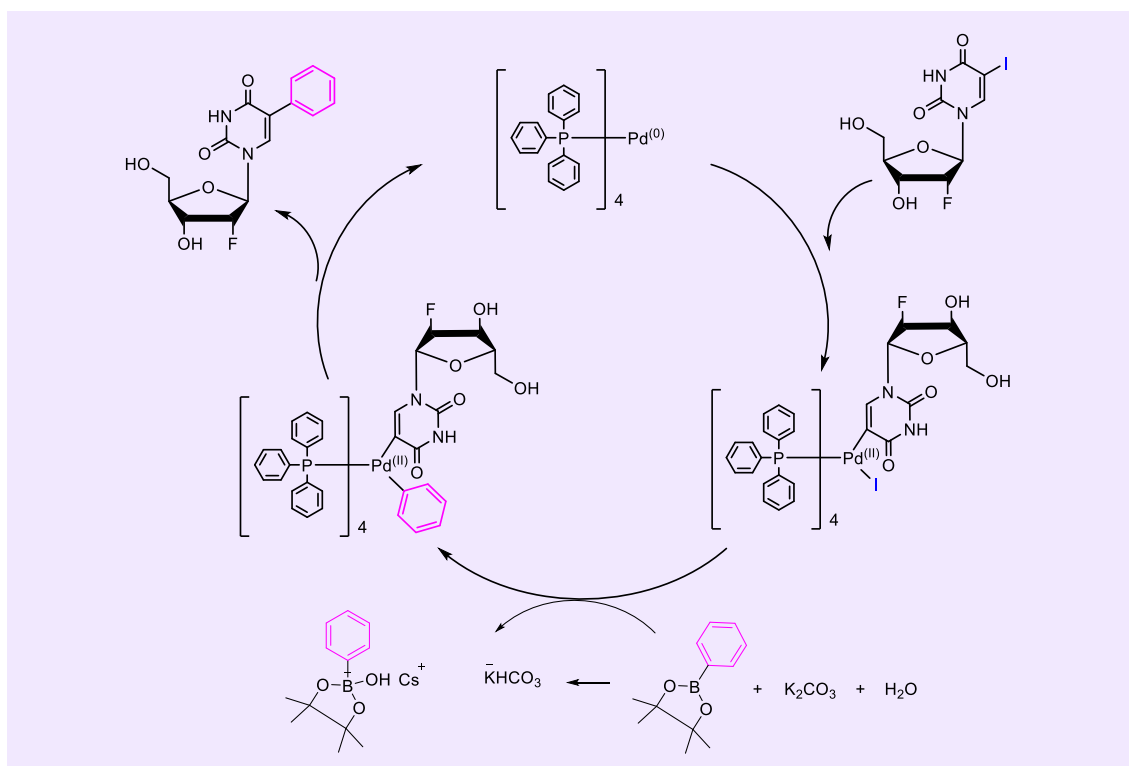


Figure 3.5. Suzuki-Miyaura reaction mechanism scheme.[383]

Palladium(II) acetate 'Pd(OAc)$_2$' is often used with triphenylphosphine to create an active catalyst species.[384] This is done by converting Pd(0) to Pd(II) (figure 3.6).[385] Amatore and Jutand have demonstrated that phosphine-induced reduction of 'Pd(OAc)$_2$' goes via trans-[Pd(OAc)$_2$(PPh3)$_2$]. These intermediate species have been observed by ^{31}P NMR spectroscopic studies. The results showed that Pd$^0$(PPh$_3$)$_n$ were created from these reaction.[386] Studies leading

on from this showed that increasing the Pd/PPh$_3$ ratio to 1:3 lead to the formation of Pd$^0$(PPh$_3$)$_n$(OAc) with the side product being O=PPh$_3$.[387] Pd$^0$(PPh$_3$)$_2$ and Pd$^0$(PPh$_3$)$_2$(OAc) both react with organohalides by oxidative addition (figure 3.6).[383]
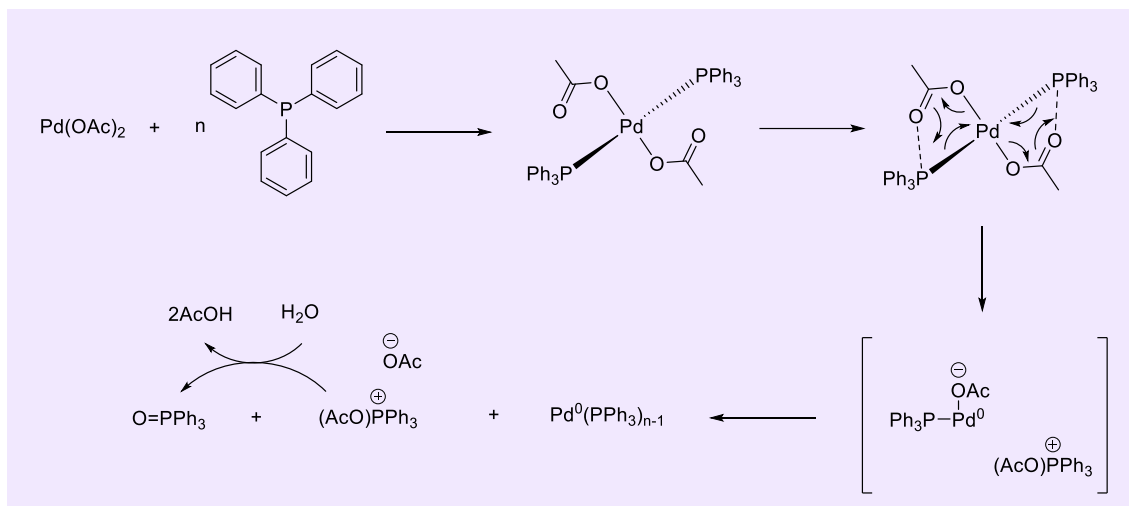


Figure 3.6. Reaction mechanism scheme of palladium acetate using triphenylphosphine.[383]

### 3.1.2.2. Stille Coupling Reaction

A Stille cross coupling reaction using tetravinyltin along with tetrakis(triphenylphosphine)palladium(0) as the catalyst can be used to create a vinyl modification. A Stille reaction is a cross coupling reaction of organic electrophiles with organostannanes. This was first done by Eaborn[388] and Kosugi-Migita.[389] Stille then developed the mechanism and catalytic cycle causing it to be called the Stille[390] reaction. The advantages of the Stille reaction are that it is compatible with a large variety of functional groups and it is mild process. These advantages mean it is often used in the synthesis of highly complex molecules.[391] Organostannanes are not very sensitive to moisture and oxygen, which allows for harsher reaction conditions. Sn contamination can still be a concern for this reaction, especially when it is used to produce pharmaceuticals. This is down to the toxicity of Sn, the commonly used derivatives of tri-*n*-butyltin have a LD$_{50}$ in the range of 100-300 mg kg$^{-1}$.[392] Toxicologists require a upper limit of Sn to be ~20 ppm.[393]

Pfizer used a stille reaction in one of the steps to synthesise [2-(3- Methyl-3H-imidazol-4-yl)-thieno[3,2-b]pyridin-7-yl]-(2-methyl1H-indol-5-yl)-amine (figure 3.7). When looking to do a large scale preparation of imidazolethienopyridine for use in pharmaceutical processes, their research shows that a Stille reaction was the most suitable for this.[394] Other coupling routes were trailed however the Stille was most suited because it provides a robust method that can be easily up scaled.
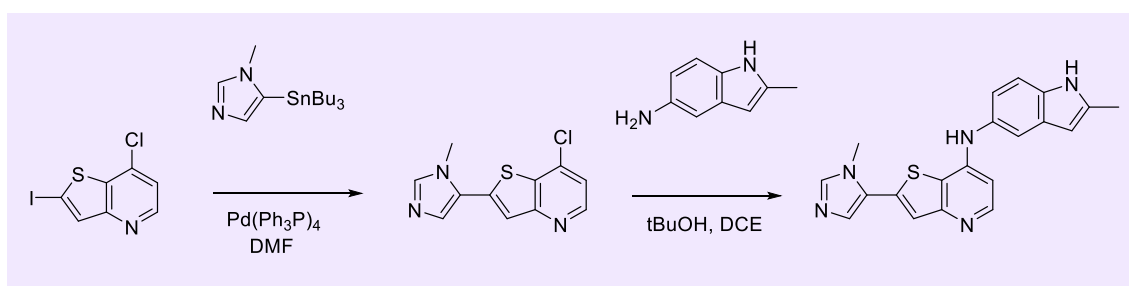


Figure 3.7. Reaction mechanism scheme of the synthesis of [2-(3- methyl-3H-imidazol-4-yl)-thieno[3,2-b]pyridin-7-yl]-(2-methyl1H-indol-5-yl)-amine.[394]

In a standard Stille reaction there are 3 main steps: oxidative addition, transmetalation and reduction elimination. This is the same for most Palladium crossing coupling reactions.

1. Oxidative addition

   This step works well with aryl iodide and Pd[0] complexes with good donor ligands. Bromides do not work as well, and chlorides normally fail. Complexes with bulky donor ligands such as phosphines are more reactive.[395]

2. Transmetalation

   This mechanism requires a good bridging anionic ligand such as a halide and a ligand that leaves easily, which will help transmetalation move faster.[396] Oxidation of Pd(PPh$_3$)$_4$ can cause a phosphine to be free which can be damaging, this can be fixed with a CuX salt.[397]

3. Reductive Elimination

Reductive elimination occurs through a three-membered transition state. The palladium catalyst needs to have two groups involved, and for them to both be cis. This requires isomerisation of the products from transmetalation who have trans stereochemistry. The couplings activation energy depends on the group being coupled. More electron rich carbons require higher activation energies.[398] The reaction mechanism can be found in figures 3.7 and 3.8.
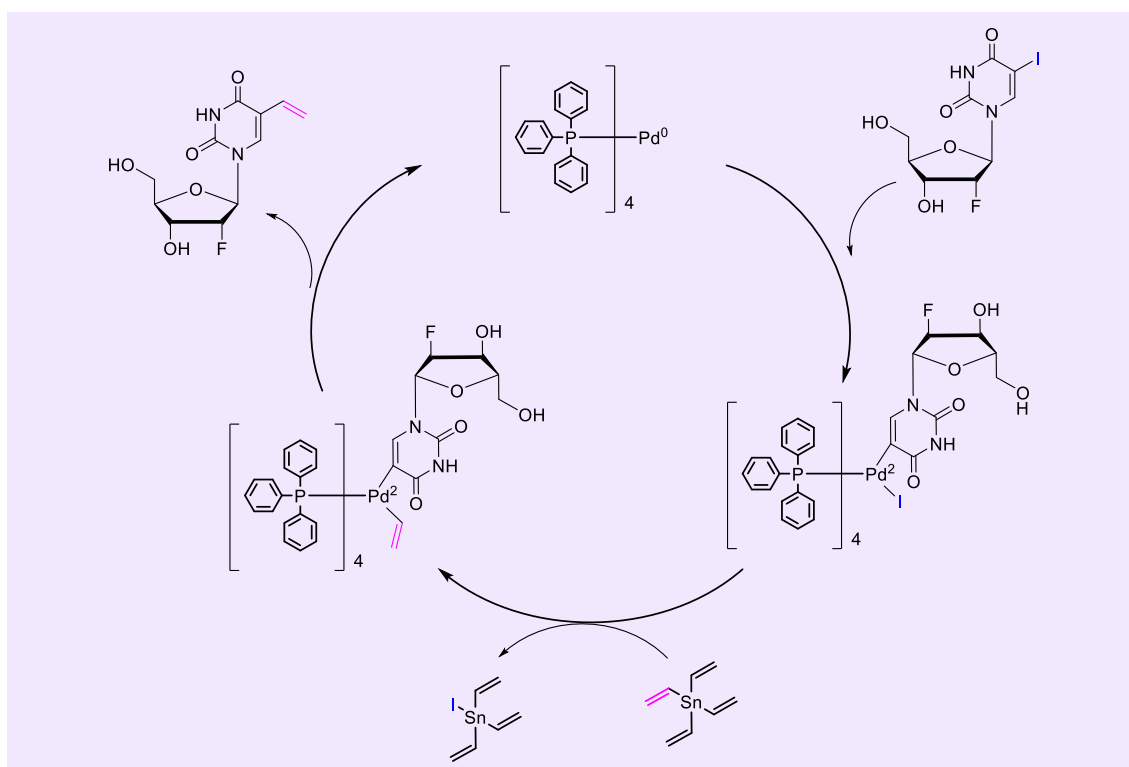


Figure 3.8. Stille Coupling Reaction Mechanism Scheme.

## 3.1.3. Chapter Aims

The aim of this chapter is to design, synthesis and analyse the modified urdine phosphoramidite compounds that will give suitable diversity such that it will effectively modulate the aptamer affinity. These compounds will be synthesised ready to be used in the aptamer library synthesis.

## 3.2. Materials & Instrumentation

### 3.2.1. Materials

2'-fluoro-2'-deoxyuridine, 4,5-dicyanoimadazole, acetic anhydride, ammonia methanol, ammonium cerium (IV)nitrate, caesium carbonate, dioxane, nitrosonium tetrafluoroborate, tetrakistriphenylphophine palladium (0) and tetravinyl tin were purchased from Acros Organic. 4-(aminomethyl)phenylboronic acid pinacol ester, 4-carboxyphenylboronic acid pinacol ester, ammonia solution, sodium bicarbonate and sodium sulfate were purchased from Alfa Aesar. Dichloromethane, diethyl Ether, dimethylformide, ethanol, ethyl acetate, magnesium sulfate, methanol, pyridine, sodium chloride and triethylamine were purchased from Fisher Scientific International. 1,1'bis(diphenylphosphineo)ferrocenepalladium(II) dichloride, 4,4'-dimethoxytrityl chloride, diisoproplethylamine, (4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl)methanamine hydrochloride, 4-dimethylaminopyridine and palladium acetate were purchased from Fluorochem. 2-cyanoethyl-diisopropl-chlorophosphoramidite was purchased from Link Technologies. Fluorenylmethoxycarbonyl, iodine, phenylboric acid pinacol ester and triphenylphosphine were purchased from Sigma Aldrich/MERCK.

All NMR samples were run in DMSO-$d_6$ purchased from Goss Scientific

### 3.2.2. Instrumentation

#### 3.2.2.1. Expedite™ 8909 DNA Synthesiser

Specification can be found in Chapter 2 section 2.2.2.1 Expedite™ 8909 DNA Synthesiser

#### 3.2.2.2. NMR

NMR spectra were obtained on a Bruker AVII 400 MHz spectrometer and were calibrated to the centre of the set solvent peak and chemical shifts were reported in parts per million (ppm).

Electrospray mass spectra were recorded on a Bruker micrOTOF-Q II mass spectrometer. The samples were analysed by injecting 2 μL of 0.1 mg/mL solutions into a flowing stream of 95% methanol, 10 mM ammonium acetate at a flow rate of 20 μL/min. The samples were analysed in both positive and negative ion modes. All the masses are mono isotopic and lock mass corrected and therefore should be within 5 ppm of the calculated mass.

# 3.3. Experimental Methods

### 3.3.1. Step 1: Acetyl protection of 2'-fluoro-2'-deoxyuridine (3.1)[399]

2'-fluoro-2'-deoxyuridine (3.1) (3.00 g, 12.19 mmol) was dissolved in pyridine (150.00 mL) and then cooled to 0°C. Acetic anhydride (12.39 g, 121.36 mmol) was added, and the reaction was stirred for 1 hour at 0°C, then for an additional hour at room temperature. The reaction was monitored by TLC (ethyl acetate/MeOH 20:1). The solvent was then removed in vacuo. The residue was then dissolved in methanol (100.00 mL) and the solvent was removed *in vacuo* again. A white powdered solid was collected, 2'-Fluoro-3',5'-di-O-acetyl-2'-deoxyuridine (3.2) (3.82 g, 98.7%).

$^1$H NMR (DMSO-$d_6$): $\delta$ 11.47 (s, 1H); 7.71-7.73 (d, 1H); 5.90 (dd, 1H); 5.68-79 (dd, 1H); 5.62 (ddd, 1H); 5.24-5.48 (ddd, 1H); 4.33 (dd, 1H); 4.24 (ddd, 1H); 4.12 (dd,1H); 2.12 (s, 3H); 2.04 (s, 3H). MS: m/z: calcd for $C_{13}H_{15}FN_2O_7$ ([M + Na]$^+$): 353.09, found 352.70

### 3.3.2. Step 2: Iodination of 2'-fluoro-3',5'-di-O-acetyl-2'-deoxyuridine (3.2)[399]

3.2 (1.50 g) was dissolved in acetonitrile (100.00 mL). After addition of iodine (1.14 g, 8.98 mmol) and (NH$_4$)$_2$[Ce(NO$_3$)$_6$] (2.49 g, 5.32 mmol) the reaction was heated to 95°C and stirred for 3hours under reflux. The reaction was monitored by TLC (ethyl acetate/MeOH 20:1). The

solvent was removed *in vacuo* and a silica column was run (Ethyl Acetate/MeOH 20:1) to product an orange oil 3',5'-di-O-acetyl-2'-desoxy-2'-fluoro-5-iodo-uridine (**3.3**) (1.5 g, 73.04%).

^{1}H NMR (DMSO-d$_6$): $\delta$ 11.87 (s, 1H); 8.20 (s, 1H); 5.84-90 (dd, 1H); 5.46-61 (ddd, 1H); 5.23-31 (ddd, 1H); 4.33-37 (dd, 1H); 4.27-31 (ddd, 1H); 4.17-21 (dd,1H); 2.11 (s, 3H); 2.08 (s, 3H). MS: m/z: calcd for C$_{13}$H$_{14}$FIN$_2$O$_7$ ([M + H]$^+$): 456.98, found 456.9 ([M + Na]$^+$): 479.98, found 479.8.

### 3.3.3. Step 3U-Ph: Addition of functional group U-Ph to 3',5'-di-O-acetyl-2'-desoxy-2'-fluoro-5-iodo-uridine (3.3) using Suzuki-Miyaura cross coupling reaction[400]

To a solution of **3.3** (1.51 g), Cs$_2$CO$_3$ (2.39 g, 7.35 mmol), phenylboronic acid pinacol ester (0.63 g, 3.11 mmol) and PPh$_3$ (0.302 g, 1.15 mmol) in dioxane (50.00 mL) and water (4.50 mL) was added to Pd(OAc)$_2$ (0.63 g, 2.82 mmol) at 25°C under N$_2$ current. The mixture was refluxed at 90°C and stirred for 16 hours. The reaction was monitored by TLC (ethyl acetate/MeOH = 20:1). The black palladium precipitate was filtered off. The reaction mixture was diluted with water (10.00 mL) and extracted with ethyl acetate (20.00 mL x 2). The organic layers were then washed with saturated aqueous NaCl (10 mL), dried over MgSO$_4$ and filtered. This was then concentrated under reduced pressure to give a residue. 3',5'-di-O-acetyl-2'-desoxy-2'-fluoro-5-phenyl-uridine was produced as a sticky yellow oil (**3.4**) (1.53 g, 87.5%).

^{1}H NMR (DMSO-d$_6$): $\delta$ 11.71 (s, 1H); 7.94 (s, 1H); 7.32-7.74 (m, 5H); 5.86-6.01 (dd, 1H); 5.48-69 (ddd, 1H); 5.28-36 (ddd, 1H); 4.33-38 (dd, 1H); 4.26-31 (ddd, 1H); 4.16-21 (dd,1H); 2.13 (s, 3H); 2.08 (s, 3H). MS: m/z: calcd for C$_{19}$H$_{19}$FN$_2$O$_7$ ([M + Na]$^+$): 429.12, found 429.00

### 3.3.4. Step 3U-Vi: Addition of functional group U-Vi to 3',5'-di-O-acetyl-2'-desoxy-2'-fluoro-5-iodo-uridine (3.3) using Suzuki-Miyaura cross coupling reaction[400]

To a solution of **3.3** (0.10 g), Cs$_2$CO$_3$ (0.16 g), Trans-1-propenylboronic acid pinacol ester (0.04 g) and PPh$_3$ (0.02 g) in dioxane (50.00 mL) and water (4.50 mL) was added to Pd(OAc)$_2$ (0.04 g) at 25°C under N$_2$ current. The mixture was refluxed at 90°C and stirred for 16 hours. The reaction

was monitored by TLC (ethyl acetate/MeOH = 20:1). The black palladium precipitate was filtered off. The reaction mixture was diluted with water (10.00 mL) and extracted with ethyl acetate (20 mL x 2). The organic layers were then washed with saturated aqueous NaCl (10.00 mL), dried over MgSO$_4$ and filtered. This was then concentrated under reduced pressure to give a residue. The desired product was not achieved.

### 3.3.5. Step 3U-Vi.2: Addition of functional group U-Vi to 3',5'-di-O-acetyl-2'-desoxy-2'-fluoro-5-iodo-uridine (3.3) using Stille cross coupling reaction[401]

**3.3** (1.00 g), tetravinyltin (0.49 ml, 1.15 mmol) and tetrakistriphenylphosphinepalladium(0) (0.19 g, 0.16 mmol) in 15.00 mL of DMF was stirred under nitrogen at 75°C for 17hours under N$_2$ current. The reaction was monitored by TLC (DCM/MeOH = 10:1). The black palladium precipitate was filtered off. The reaction mixture was diluted with water (25.00 mL x 2) and extracted with diethyl ether (25.00 mL x 2). The diethyl layers were combined and dried over MgSO$_4$, filtered, and concentrated under reduced pressure to give a residue. 3',5'-di-O-acetyl-2'-desoxy-2'-fluoro-5-vinyl-uridine was produced as a yellow oil (**3.6**) (0.37 g, 35.0%).

^{1}H NMR (DMSO-d$_6$): $\delta$ 11.62 (s, 1H); 7.88 (s, 1H); 6.33-41 (dd,1H); 5.85-99 (q, 1H); 5.50-65 (dd, 1H) 5.25-5.33 (m, 2H); 5.17-5.19 (dd, 1H); 4.33-37 (dd, 1H); 4.28-31 (ddd, 1H); 4.17-22 (dd,1H); 2.11 (s, 3H); 2.03 (s, 3H). MS: m/z: calcd for C$_{15}$H$_{17}$FN$_2$O$_7$ ([M + Na]$^+$): 379.10, found 379.0

### 3.3.6. Step 3U-NH$_2$.1: Addition of functional group U-I to 3',5'-di-O-acetyl-2'-desoxy-2'-fluoro-5-iodo-uridine (3.3) using Suzuki-Miyaura cross coupling reaction[400]

To a solution of **3.3** (0.10 g), Cs$_2$CO$_3$ (0.16 g), 4-(aminomethyl)phenylboronic acid pinacol ester (0.056 g) and PPh$_3$ (0.02 g) in dioxane (5.00 mL) and water (4.50 mL) was added to Pd(OAc)$_2$ (0.04 g) at 25°C under N$_2$ current. The mixture was heated to 90°C and stirred for 16 hours. The reaction mixture was diluted with water (10.00 mL) and extracted with ethyl acetate (20.00 mL x 2). The organic layers were then washed with saturated aqueous NaCl (10.00 mL), dried over

MgSO$_4$, filtered, and concentrated under reduced pressure to give a residue. $^1$H NMR, $^{13}$C NMR and mass spectrometry were run on the product. The desired product was not achieved.

### 3.3.7. Step 3U-NH$_2$.2: Synthesis of Fmoc protected 4-(aminomethyl)phenylboronic acid pinacol ester

Fmoc (0.11 g) and 4-(aminomethyl)phenylboronic acid pinacol ester (0.10 g) were dissolved in 5.00 mL of ethanol. This solution was sonicated for 20 minutes at room temperature. $^1$H NMR was run on the product, Fmoc-4-(aminomethyl)phenylboronic acid pinacol ester (**3.8**).

$^1$H NMR (DMSO-d$_6$): $\delta$ 8.60 (s, 1H); 7.85-87 (d, 2H); 7.67-71 (q, 2H); 7.50-52 (d, 2H); 7.36-37 (d, 2H); 7.30-31 (d, 2H); 4.55 (s, 2H); 4.14 (s, 1H); 4.02-04 (d, 2H)

### 3.3.8. Step 3U-NH$_2$.3: Addition of functional group U-NH$_2$ to 3',5'-di-O-acetyl-2'-desoxy-2'-fluoro-5-iodo-uridine (3.3) using Suzuki-Miyaura cross coupling reaction[400]

To a solution of **3.3** (0.12 g), Cs$_2$CO$_3$ (0.17 g), **3.8** (0.23 g in excess) and PPh$_3$ (0.02 g) in dioxane (100.00 mL) and water (9.00 mL) was added to Pd(OAc)$_2$ (0.05 g) at 25°C under N$_2$ current. The mixture was heated to 90°C and stirred for 16 hours. The reaction was monitored by TLC (ethyl acetate/MeOH = 20:1). The reaction mixture was diluted with water (10.00 mL) and extracted with ethyl acetate (20.00 mL x 2). The organic layers were then washed with saturated aqueous NaCl (10.00 mL), dried over MgSO$_4$, filtered, and concentrated under reduced pressure to give a residue. $^1$H NMR and mass spectrometry were run on the product. The desired product was not achieved.

### 3.3.9 Step 3U-NH$_2$.4: Addition of functional group U-NH$_2$ to 3',5'-di-O-acetyl-2'-desoxy-2'-fluoro-5-iodo-uridine (3.3) using Suzuki-Miyaura cross coupling reaction

Method 1[400]: To a solution of **3.3** (0.12 g), Cs$_2$CO$_3$ (0.17 g), 4-Carboxylphenylboronic acid pinacol ester (0.05 g) and PPh$_3$ (0.02 g) in dioxane (100.00 mL) and water (9.00 mL) was added to

Pd(OAc)$_2$ (0.05 g) at 25°C under N$_2$ current. The mixture was heated to 90°C and stirred for 16 hours. The reaction was monitored by TLC (ethyl acetate/MeOH = 20:1). The reaction mixture was diluted with water (10.00 mL) and extracted with ethyl acetate (20.00 mL x 2). The organic layers were then washed with saturated aqueous NaCl (10 mL), dried over MgSO$_4$, filtered and concentrated under reduced pressure to give a residue. The desired product was not achieved.

Method 2[402]: To a solution of **3.3** (0.05 g), Cs$_2$CO$_3$ (0.07 g) and 4-Carboxylphenylboronic acid pinacol ester (0.01 g) in dioxane (50.00 mL) and water (4.5.00 mL) was added to 1,1'-bis(diphenylphosphino)ferrocenepalladium (II) dichloride (0.07 g) at 25°C under N$_2$ current. The mixture was heated to 90°C and stirred for 16 hours overnight. The reaction was monitored by TLC (ethyl acetate/MeOH = 20:1). The reaction mixture was diluted with water (10 mL) and extracted with ethyl acetate (20.00 mL x 2). The organic layers were then washed with sat. aqueous NaCl (10.00 mL), dried over MgSO$_4$, filtered, and concentrated under reduced pressure to give a residue. $^1$H NMR and mass spectrometry were run on the product. The desired product was not achieved.

### 3.3.10 Step 3U-NH$_2$.5: Addition of functional group U-NH$_2$ to 3',5'-di-O-acetyl-2'-desoxy-2'-fluoro-5-iodo-uridine (**3.3**)[403]

**3.3** (0.05 g, 1 mmol) was dissolved in 50.00 mL of DMF and then treated with nitrosonium tetrafluoroborate (0.29 g, 5.70 mmol) for 24 hours. The reaction was terminated by the addition of water (1.00 mL). The reaction mixture was diluted with ethyl acetate (50.00 mL) and washed with sodium bicarbonate. The organic layer was separated and dried over sodium sulfate. Crude product was purified on silica gel column using ethyl acetate/hexane (1:1). $^1$H NMR and mass spectrometry were run on the product. $^1$H NMR and mass spectrometry were run on the product. The desired product was not achieved.

### 3.3.11. Step 3U-NH₂.6: Addition of functional group U-NH₂ to 3',5'-di-O-acetyl-2'-desoxy-2'-fluoro-5-iodo-uridine (3.3)[404]

**3.3** (0.04 g) was treated with 15.00 mL of ammonia solution (35%) at 55°c, in a sealed screw-cap bottle. The solution was cooled and concentrated under reduced pressure. $^1$H NMR and mass spectrometry were run on the product. The desired product was not achieved.

### 3.3.12. Step 4U-Ph: Removal of acetyl protecting groups from 3',5'-di-O-acetyl-2'-desoxy-2'-fluoro-5-phenyl-uridine (3.4)

**3.4** (1.53 g) was dissolved in an NH₃/MeOH solution (7N, 15.00 mL) and stirred at room temperature for 20 hours. The reaction was monitored by TLC (ethyl acetate /MeOH 20:1). The solvent was removed in vacuo and a solid of 2'-desoxy-2'-fluoro-5-phenyl-uridine was collected (**3.15**) (0.98 g, 81.1%).

$^1$H NMR (DMSO-d₆): δ 11.71 (s, 1H); 7.94 (s, 1H); 7.32-7.74 (m, 5H); 5.86-6.01 (dd, 1H); 5.48-69 (ddd, 1H); 5.28-36 (ddd, 1H); 4.33-38 (dd, 1H); 4.26-31 (ddd, 1H); 4.16-21 (dd,1H). MS: m/z: calcd for $C_{19}H_{19}FN_2O_7$ ([M + Na]⁺): 345.10, found 345.0

### 3.3.13. Step 4U-Vi: Removal of acetyl protecting groups from 3',5'-di-O-acetyl-2'-desoxy-2'-fluoro-5-vinyl-uridine (3.6)

**3.6** (0.37 g) was dissolved in an NH₃/MeOH solution (7N, 15.00 mL) and stirred at room temperature for 20 hours. The reaction was monitored by TLC (ethyl acetate /MeOH 20:1). The solvent was removed in vacuo and a solid of 2'-desoxy-2'-fluoro-5-vinyl-uridine was collected (**3.16**) (0.25 g, 92.6%).

$^1$H NMR (DMSO-d₆): δ 11.53 (s, 1H); 8.28 (s, 1H); 6.36-39 (d, 1H); 6.30-32 (d, 1H); 5.92-94 (t, 1H); 5.28-36 (t, 1H); 5.42-65 (dd, 1H); 5.11-14 (q, 1H); 4.99 (s,1H); 4.20-30 (dd,1H); 3.82-89 (t,1H); 3.64-74 (t,1H); 3.61 (s,1H). MS: m/z: calcd for $C_{11}H_{13}FN_2O_5$ ([M + H]⁺): 272.08, found 272.91.

### 3.3.14. Step 4U-I: Removal of acetyl protecting groups from 3',5'-di-O-acetyl-2'-desoxy-2'-fluoro-5-iodo-uridine (3.3)

**3.3** (1.53 g) was dissolved in an $NH_3$/MeOH solution (7N, 15.00 mL) and stirred at room temperature for 20 hours. The reaction was monitored by TLC (ethyl acetate/MeOH 20:1). The solvent was removed in vacuo and a solid of 2'-desoxy-2'-fluoro-5-iodo-uridine was collected (**3.17**) (1.21 g, 96.8%).

$^1$H NMR (DMSO-$d_6$): $\delta$ 11.75 (s, 1H); 8.55 (s, 1H); 5.83-5.87 (d, 1H); 5.63-5.64 (t, 1H); 5.43 (s, 1H); 4.97-5.11 (dd, 1H); 4.14-4.18 (dd, 1H); 3.88-3.90 (t, 1H); 3.81-3.83 (t,1H); 3.58(s,1H). MS: m/z: calcd for $C_9H_{10}FN_2O_5$ ([M + H]$^+$): 372.96, found 373.90. ([M + Na]$^+$): 395.10, found 394.85.

### 3.3.15. Step 5U-Ph: Addition of 4,4'-dimethoxytrityl protecting group to 2'-desoxy-2'-fluoro-5-phenyl-uridine (3.15)

**3.15** (0.06 g) and DMAP (0.0013 g, 0.0105 mmol) were dissolved in pyridine (10.00 mL). 4,4'-Dimethoxytrityl chloride (0.077 g, 0.23 mmol) was added in portions over 1 hour dropwise at 0°C. After complete consumption of the starting material (18 hours), MeOH (2.00 mL) was added, and the resulting mixture was concentrated under vacuum. The product was purified by column chromatography using a basic alumina stationary phase and eluted with DCM/MeOH (100:1) to produce a yellow oil (**3.18**) (0.14 g, 186.0%).

$^1$H NMR (DMSO-$d_6$): $\delta$ 9.89 (s, 1H); 8.19 (s, 1H); 7.27-7.31 (m, 3H); 7.19-7.23 (m, 9H); 7.06-7.10 (m, 2H) 6.84-6.89 (m, 4H); 6.22 (s, 1H); 3.91 (q, 1H); 3.87 (t, 1H); 43.66 (s, 1H); 3.55 (d, 1H); 3.07-3.07 (dd,2H). MS: m/z: calcd for $C_{36}H_{33}FN_2O_7$ ([M + Na]$^+$): 647.23, found 647.20

### 3.3.16. Step 5U-Vi: Addition of 4,4'-dimethoxytrityl protecting group to 2'-desoxy-2'-fluoro-5-vinyl-uridine (3.16)

**3.16** (0.07 g) and DMAP (0.0013 g, 0.011 mmol) were dissolved in pyridine (10.00 mL). DMTCl (0.13 g, 0.40 mmol) was added in portions over 1 hour dropwise at 0°c. After complete

consumption of the starting material (18 hours), MeOH (2.00 mL) was added, and the resulting mixture was concentrated under vacuum. The product was purified by column chromatography using a basic alumina stationary phase and eluted with DCM/MeOH (100:1) to produce a yellow oil (**3.19**) (0.20 g).

$^1$H NMR (DMSO-d$_6$): $\delta$ 11.56 (s, 1H); 8.31 (s, 1H); 7.25-7.36 (m, 11H); 6.86-6.88 (m, 4H); 6.36-6.41 (d, 1H); 6.30-6.33 (t, 1H); 5.72-5.96 (dd, 2H); 5.49-5.50 (q, 1H); 5.12-5.13 (t,1H); 5.01 (s,1H); 4.12-4.22 (dd,1H); 3.86-3.92 (dd, 2H); 3.72 (s, 6H). MS: m/z: calcd for C$_{32}$H$_{31}$FN$_2$O$_7$ ([M + H]$^+$): 574.21 , found 574.00.

### 3.3.17. Step 5U-I: Addition of 4,4'-dimethoxytrityl protecting group to 2'-desoxy-2'-fluoro-5-iodo-uridine (3.17)

**3.17** (0.12 g) and DMAP (0.02 g, 0.02 mmol) were dissolved in pyridine (10.00 mL). 4,4'-Dimethoxytrityl chloride (0.24 g, 0.70 mmol) was added in portions over 1 hour dropwise at 0$^o$C. After complete consumption of the starting material (18 hours), MeOH (2.00 mL) was added and the resulting mixture was concentrated under vacuum. The product was purified by column chromatography using a basic alumina stationary phase and eluted with DCM/MeOH (100:1) to produce a yellow oil was (**3.20**) (0.333 g).

$^1$H NMR (DMSO-d$_6$): $\delta$ 11.79 (s, 1H); 8.58 (d, 1H); 7.08-7.37 (m, 11H); 6.84-6.91 (m,4H); 5.83-5.90 (d, 1H); 5.45-5.67 (dd,1H); 4.99-5.24 (t, 1H); 4.38 (s, 1H); 4.25-4.35 (dd, 1H); 3.93 (t, 1H); 3.74 (s, 6H); 3.64 (d, 1H). MS: m/z: calcd for C$_{30}$H$_{28}$FN$_2$O$_7$ ([M + H]$^+$): 697.09 , found 697.00.

### 3.3.18. Step 5U: Addition of 4,4'-dimethoxytrityl protecting group to 2'-desoxy-2'-fluoro-uridine (3.1)

**3.1** (0.12 g) and DMAP (0.002 g, 0.02 mmol) were dissolved in DMF (10.00 mL) with a few drops of trimethylamine. 4,4'-Dimethoxytrityl chloride (0.50 g, 1.48 mmol) was added in portions over 1 hour dropwise at 0$^o$C. After complete consumption of the starting material (18 hours), MeOH

(2.00 mL) was added, and the resulting mixture was concentrated under vacuum. The product was purified by column chromatography using a basic alumina stationary phase and eluted with DCM/MeOH (100:1) to produce a yellow oil (**3.24**) (0.10 g).

$^1$H NMR (DMSO-d$_6$): $\delta$ 11.32 (s, 1H); 7.43 (d, 1H); 6.89-7.30 (m, 13H); 6.55 (d, 1H); 5.47 (d,1H); 4.60-4.61 (q, 1H); 4.55 (t, 1H); 4.37 (s, 1H); 4.30 (t, 1H); 3.81 (d, 6H); 3.36 -3.61 (dd, 2H).

### 3.3.19. Step 6U-Ph: Addition of 2-cyanoethyl diisopropyl chlorophosphoramidite to DMT protected 2'-desoxy-2'-fluoro-5-vinyl-uridine (3.18)

**3.19** (0.31 g), and dimethylaminopyridine (0.05 g, 0.41 mmol) were dissolved in 50.00 mL of dry DCM, diisopropylethylamine (0.30 mL, 2.33 mmol) and 2-cyanoethyl-diisopropyl-chlorophosphoramidite (0.30 mL, 1.25 mmol) were added. Under an atmosphere of nitrogen, the mixture was then stirred at room temperature for 2 hours. After the DCM was removed *in vacuo* a thick yellow oil was obtained (**3.21**) (0.60 g). $^{31}$P NMR was run on the product, to prove the addition of the phosphoramidite group. No other analysis was run on it because of the air sensitive nature of the product.

$^{31}$P NMR (DMSO-d$_6$ capillary in DCM): $\delta$ 150.00 (s, 1P) (0.77:0.23)

### 3.3.20. Step 6U-Vi: Addition of 2-cyanoethyl diisopropyl chlorophosphoramidite to DMT protected 2'-desoxy-2'-fluoro-5-vinyl-uridine (3.19)

**3.19** (0.88 g), and dimethylaminopyridine (0.14 g, 1.14 mmol) were dissolved in 50.00 mL of dry DCM, diisopropylethylamine (0.84 mL, 6.51 mmol) and 2-cyanoethyl-diisopropyl-chlorophosphoramidite (0.83 mL, 3.50 mmol) were added. Under an atmosphere of nitrogen, the mixture was then stirred at room temperature for 2 hours. After the DCM was removed *in vacuo* a thick yellow oil was obtained (**3.22**) (1.63 g). $^{31}$P NMR was run on the product, to prove the addition of the phosphoramidite group. No other analysis was run on it because of the air sensitive nature of the product.

$^{31}$P NMR (DMSO-d$_6$ capillary in DCM): $\delta$ 139.00 (s, 1P)

### 3.3.21. Step 6U-I: Addition of 2-cyanoethyl diisopropyl chlorophosphoramidite to DMT protected 2'-desoxy-2'-fluoro-5-iodo-uridine (3.20)

**3.23** (1.07 g), and dimethylaminopyridine (0.17 g, 1.39 mmol) was dissolved in 50.00 mL of dry DCM, diisopropylethylamine (1.03 mL, 7.95 mmol) and 2-cyanoethyl-diisopropyl-chlorophosphoramidite (1.01 mL, 4.25 mmol) were added. Under an atmosphere of nitrogen, the mixture was then stirred at room temperature for 2 hours. After the DCM was removed *in vacuo* to obtained thick yellow oil (**3.23**) (1.98 g). $^{31}$P NMR was run on the product, to prove the addition of the phosphoramidite group. No other analysis was run on it because of the air sensitive nature of the product.

$^{31}$P NMR (DMSO-d$_6$ capillary in DCM): $\delta$ 149.00 (s, 1P)

### 3.3.22. Step 6U: Addition of 2-cyanoethyl diisopropyl chlorophosphoramidite to DMT protected 2'-desoxy-2'-fluoro-uridine (3.24)

**3.24** (0.30 g), and dimethylaminopyridine (0.05 g, 0.43 mmol) were dissolved in 50.00 mL of dry DCM, diisopropylethylamine (0.29 mL, 2.02 mmol) and 2-cyanoethyl-diisopropyl-chlorophosphoramidite (0.28 mL, 1.20 mmol) were added. Under an atmosphere of nitrogen, the mixture was then stirred at room temperature for 2 hours. After the DCM was removed in vacuo a thick yellow oil was obtained (**3.25**) (0.61 g). $^{31}$P NMR was run on the product, to prove the addition of the phosphoramidite group. No other analysis was run on it because of the air sensitive nature of the product.

$^{31}$P NMR (DMSO-d$_6$ capillary in DCM): $\delta$ 150.00 (s, 1P) (0.63:0.37)

### 3.3.23. Synthesis of ARP25U

ARP25U: 5'-fUfUfUfUfUfUfUfUfUfUfUfUfUfUfUfUfUfUfUfUfUfUfUfUfU-3'

The Expedite™ 8909 DNA synthesiser was used to synthesise ARP25U on Universal UnyLinker support. **3.25** (1.00 g) was dissolved in 20.00 mL DCM and put in bottles and screwed into the synthesiser lines. The synthesis reagents were Oxidizer (0.02 M iodine, 20% pyridine), Cap A Mix (THF/pyridine/acetic anhydride 8:1:1), Cap B Mix (10% methylimidazole in THF), deblock (3% trimethylamine in DCM) and ETT activator (0.25 M, 5-ethylthio-1H-tetrazole in acetonitrile) were screwed into the synthesiser lines. A leak test was run to check nitrogen was not leaking from the lines. When passed, the lines are then flushed with the new reagents added to them. Universal UnyLinker support (0.021 g) was added to the column, which was then placed on the synthesiser, which was then flushed with acetonitrile several times. The sequence and protocol are then selected using Expedite™ 8909 software connected to the Expedite™ 8909 DNA synthesiser. The sequence was then run and monitored using the trityl monitor. The protocol for coupling was extended 10x more than a standard DNA 1 µmol protocol.

# 3.4. Results and Discussion

## 3.4.1. Modified Uridine Designs

The uridine monomers in MinE07 were chosen to be modified as they are synthetic accessibility, utilising known chemistry as well as financially available starting materials. The uridines were modified at the C(5) position because modifications at the C(5) position of pyrimidine nucleosides can enhance their bioavailability, bioactivity and biostability.[405] The synthesis of C(5)-substituted nucleotides have gained importance because of their potential medicinal applications.[406] 5-halopyrimindies exhibit anticancer and antiviral as well as photosensitising properties. These compounds were used in research that discovered key information concerning their affinity for transporter proteins and their cellular biochemistry.[407] There are several 5-halo-modified pyrimidine drugs available such as 1-hexylcarbamoyl-5-fluorouracil which is a treatment for colorectal cancer[408] and 5-fluoro-20-deoxyuridine, which is a commonly

known anticancer drug used in treatment of stomach, colorectal and kidney cancer.[409] Some other examples of 5-substituted nucleosides with confirmed biological activities are 5-ethyl-20-deoxyuridine which is an anti-HSV agent[410] and 5-ethynyl-20-deoxyuridine which is used against herpes simplex-1 strains KOS and LYONS.[411] The five modifications were designed so that they covered a range of polar, non-polar, positively charged, aromatic, double bonds, halogens and natural uridine. Molecular recognition in biological systems relies on there being specific attractive interactions between two molecules.[412] The modifications chosen could either enhance or disrupt these interactions that are already occurring between MinE07 and EGFR.

### 3.4.1.1. Aromatic Ring Addition Modification
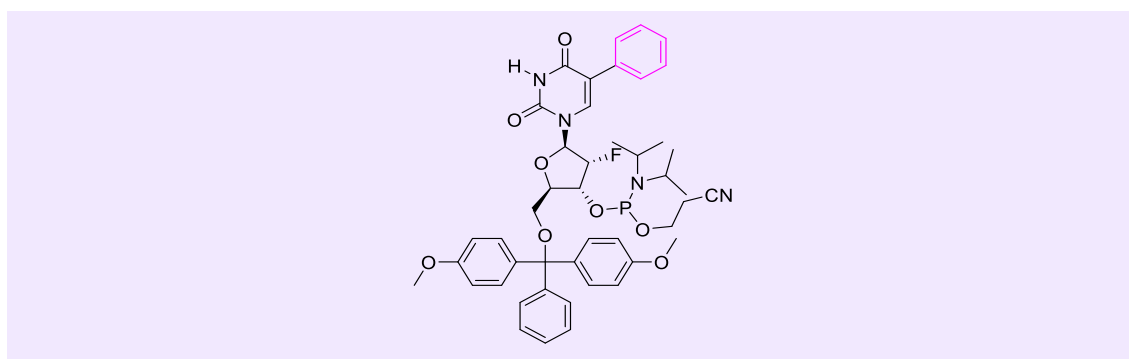


Figure 3.9. Chemical Structure of DMTr -protected 2'-desoxy-2'-fluoro-5-phenyl-uridine phosphoramidite (**3.21**).

Aromatic compounds are cyclic structures with an unsaturated hydrocarbon ring planar structure that have specific properties, which include extreme stability due to their aromaticity. They are generally non-polar, have little reactivity and are immiscible with water. Cyclic aromatic compounds undergo electrophilic substitution reactions with the ring acting as a nucleophile.[413]

There are numerous non-covalent inter- and intramolecular interactions between aromatic rings in biological and chemical processes.[414] These include self-assembly, molecular recognition, transport and catalysis.[415] Interactions between aromatic and heteroaromatic rings are key contributors to protein-ligand complexes as well as protein structures.[416] They are also

major contributors to the base pair stacking in nucleic acid structures seen in duplex DNA.[417] The π- π interactions play a very important part in RNA-protein interactions.[418]

An aromatic ring modification was chosen because of its size, the stability of the ring and the non-covalent inter- and intramolecular interactions between aromatic rings in biological and chemical processes, especially in protein ligand complexes. The addition of the aromatic ring could potentially cause more protein-ligand interactions which may enhance the affinity of MinE07 for EGFR. The size and rigidity of the ring could potentially disrupt the folding of the MinE07 aptamer, changing its shape. This could either show that the shape of the aptamer does not influence the binding affinity of MinE07 for EGFR or that the shape is hugely important for how the aptamer fits in to the receptor. This modification was also chosen as the chemistry of phenylation to a nucleotide is known (figure 3.9). Using a Suzuki-Miyaura palladium catalysed reaction has been successful.[419]

### 6.4.1.1. Aliphatic Modification

Bonding between one carbon atom and another is covalent, containing either σ or π bonds. Carbon can make single, double or triple bonds, which determines the structure generated. With one double bond its structure is trigonal planar.
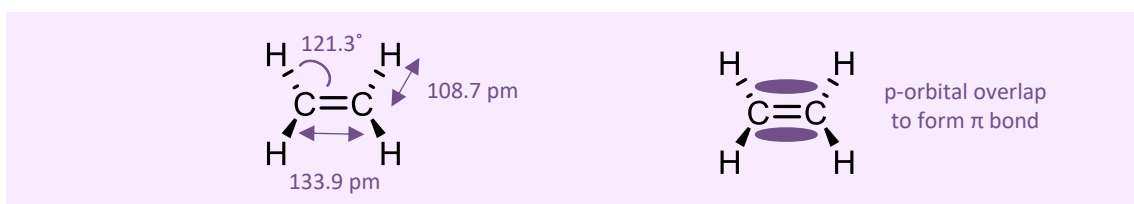


Figure 3.10. Chemical Structure of Ethylene

In ethylene each carbon atom has three $sp^2$ orbitals lying in a plane with ~120° angles and one p-orbital perpendicular to this plane (figure 3.10). When the carbon atoms approach each other, two of the $sp^2$ orbitals overlap to form a σ bond, the two p-orbitals form a π bond. There is no free rotation about the carbon-carbon σ bond because of the π bond. The C=C bond length of

133.9 pm is shorter than C-C length of 154 pm. It is also stronger with 636 kJ mol$^{-1}$ compared to 368 kJ mol$^{-1.}$ Vinyl compounds contain the hydrocarbon vinyl group CH$_2$=CH-R. A single vinyl compound can be made to polymerise forming polyvinyl compounds. These compounds can form various styrene and its derivatives. The double bond in vinyl compounds is accessible for reactions with its π electrons positioned above and below the double bond. The double bond can easily donate electrons to act like a nucleophile, so can undertake electrophilic addition. A vinyl modification was chosen because of these potential chemical interactions (figure 3.11).
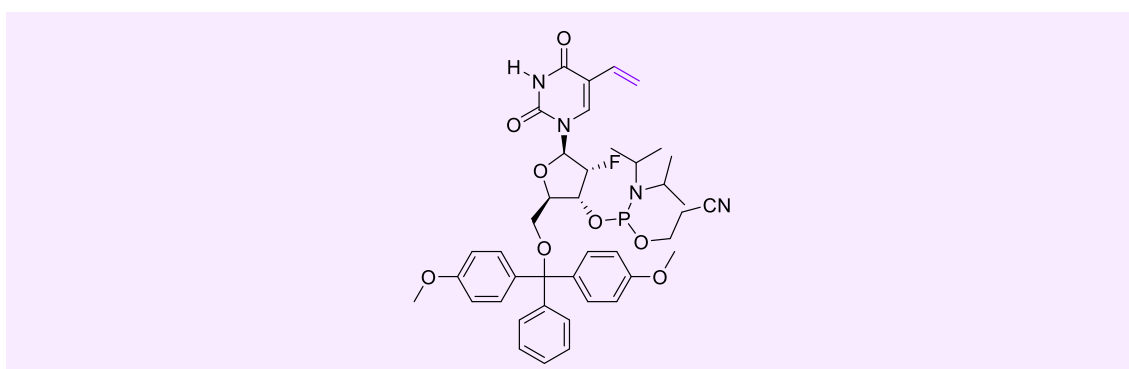


Figure 3.11. Chemical Structure of DMTr -protected 2'-desoxy-2'-fluoro-5-vinyl-uridine phosphoramidite (3.22).

### 6.4.1.2. Amino Group Modification

Hydrogen bonding is the strongest force between two molecules that has no formal charge. It is a force between a partially positive hydrogen atom and a partially negative atom such as a nitrogen or oxygen.[420] Most often hydrogen bonding occurs in molecules that contain an amino group (NH$_2$) and hydroxyl groups (OH). As the hydrogens are connected to highly electronegative atoms their electron density is drawn away from them, which leaves them with a partially positive charge and the electronegative atom having gained electron density is new partially negative. This then allows for strong electrostatic interactions between the negative atom and a hydrogen on another molecule.
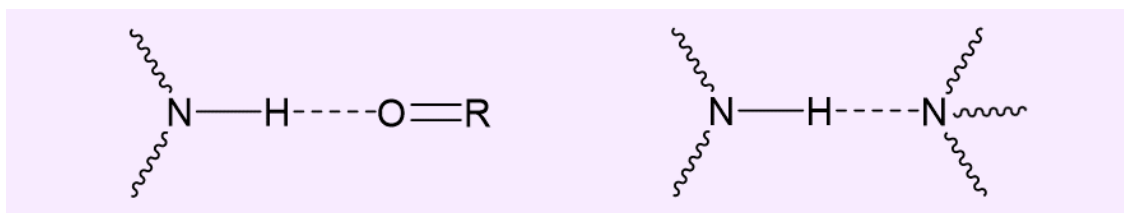
Figure 3.12. Hydrogen Bonding.

Hydrogen bonds are the most important specific interactions in biological recognition processes. They play a key role in the working of the genetic code, such as their selectivity in the purine and pyrimidine nucleic bases.[421] An amino group modification was chosen as it could potentially add or influence hydrogen bonding, mimicking base pairing we see in RNA and DNA shown in figure 3.12.
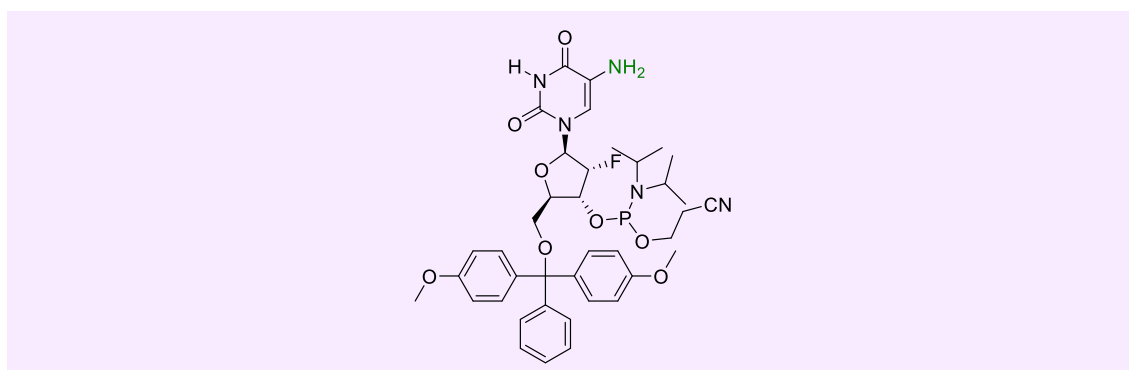


Figure 3.13: Chemical Structure of DMTr -protected 2'-desoxy-2'-fluoro-5-vinyl-uridine phosphoramidite (**3.26**).

### 3.4.1.4. Iodo Modification

Halogenated nucleic acids with either Br or I atoms are able to establish halogen bonds with properly disposed protein residues. Halogen bonding interactions are characterised as members of the "σ-hole bonding" family.[422] A σ-hole is defined as a region of positive electrostatic potential on a molecule located on the extension of a covalent bond. The σ-hole donor ability increases down the halogens F to I, meaning that halogen bonds involving iodine are the most enthalpically favoured. In medicinal chemistry, halogen atoms are commonly used in drug design and to probe molecular interactions.[413] They have been incorporated in drug

design because drug selectivity can be improved with the addition of bromine and iodine, whilst the addition of chlorine and fluorine can be used to optimise absorption, distribution, metabolism and excretion (ADME) properties and to reduce undesired reactions such as ring hydroxylation.[371] A known drug using iodine is 5-iodo-20-deoxyuridine which is used as a treatment for feline herpes virus (FHV-1). An iodination modification was chosen because halogens have proved to be very useful in medicinal chemistry because of the increased bioavailability and the improved metabolic stability the halogen provides. It is also known to improve protein-ligand interactions.[373] This modification was also chosen because the chemical synthesis is well developed (figure 3.14).[399]
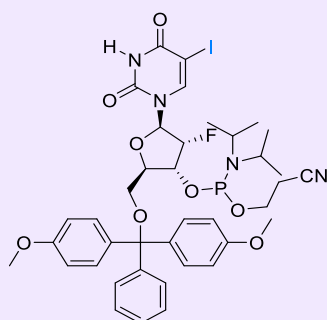


Figure 3.14. Chemical structure of DMTr -protected 2'-desoxy-2'-fluoro-5-iodo-uridine phosphoramidite (**3.23**).

### 3.4.1.5. The General Synthesis design for the Modification of Uridine

Presented below is the planned reaction scheme for the synthesis of the modified uridines (figure 3.15). R being the chosen functional groups to modify the uridine with. The acetyl protecting groups are added first to stop any unwanted reactions happening at the OH positions. These are later removed to add a DMT protecting group, which is added so the synthesis progression can be monitored in the synthesiser by the removal of it.
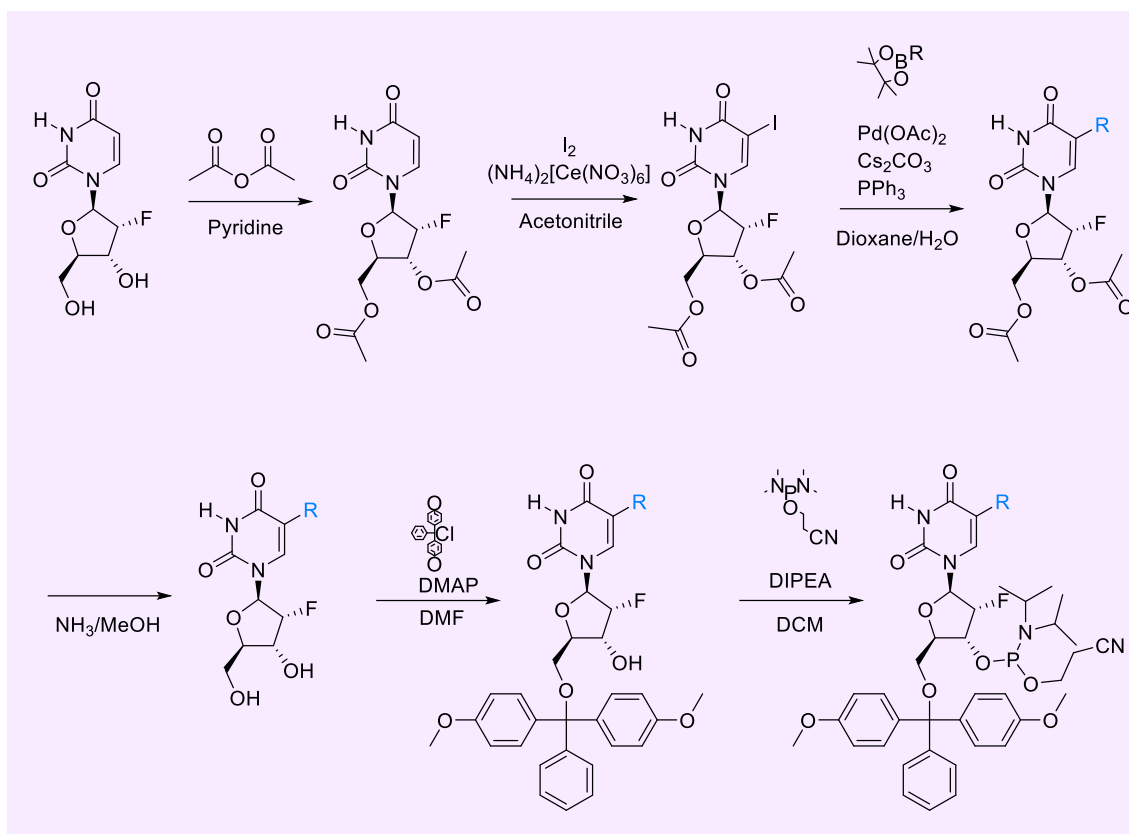
Figure 3.15. The overall planned reaction scheme for the modification of uridine.

## 3.4.2. General Modified Uridine Synthesis

### 3.4.2.1. Step 1: Acetyl protection of 2'-fluoro-2'-deoxyuridine (3.1)[399]
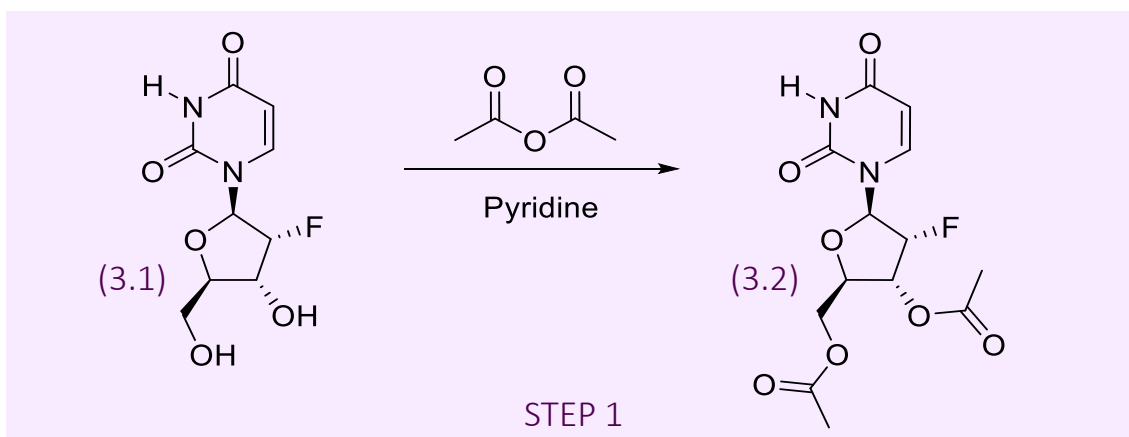


Figure 3.16. Reaction scheme for acetyl protection of 2'-Fluoro-2'-deoxyuridine (**3.1**)

Step 1 (figure 3.16) uses protecting group chemistry to add an acetyl group to both hydroxyl groups on the uridine at the C'3 and C'5 positions. This is done to stop the hydroxyl groups

152

interacting in future reactions. An acetyl group was chosen as they are very commonly used in oligonucleotide synthesis, are stable in the conditions used in future steps and are easily removed when needed. This reaction works very well and produces a high yield with minimal amounts of purification and method development needed. ^{1}H NMR spectra of the product 2'-fluoro-3',5'-di-o-acetyl-2'-deoxyuridine (**3.2**) can be found in figure 3.17. The addition of large singlet peaks at 2.05 ppm and 2.12 ppm corresponding to the new acetyl groups along with the loss of two -OH singlet peaks at 3.94 ppm and 4.37 ppm show that is reaction was successful and pure enough to carry forward to the next reaction step (some solvent peaks extra found in spectra). This was also confirmed by the ^{13}C NMR (appendix A) along with the mass spectrometry (appendix B) showing the correct mass peak plus a Na[+] ion at 352.7 m/z. The acetyl groups can be removed by treating them with a base, the most commonly used are ammonia (aqueous or gaseous) or methylamine.[423]
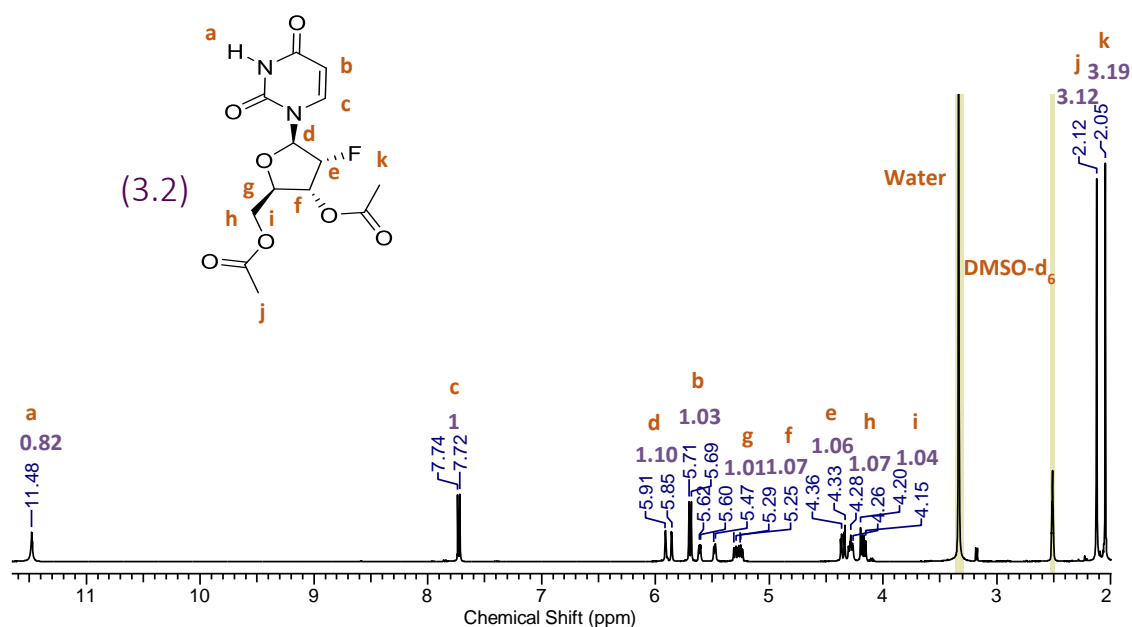


Figure 3.17. ^{1}H NMR spectra of 2'-fluoro-2'-deoxyuridine (integrations in purple above) (**3.2**). Extra peak at 3.2 ppm is caused by excess methanol from work up of the reaction.

## 3.4.2.2. Step 2: Iodination of 2'-fluoro-3',5'-di-O-acetyl-2'-deoxyuridine (**3.2**)[399]
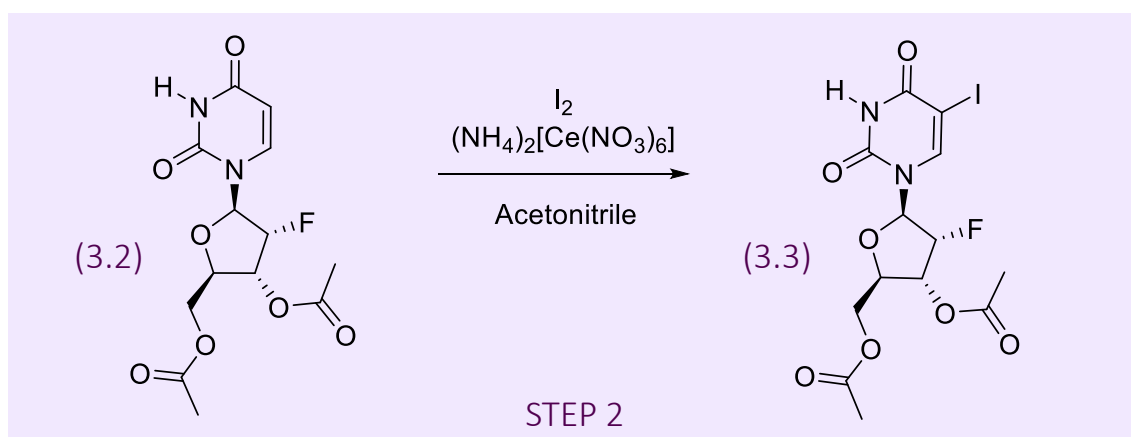


Figure 3.18. Reaction scheme for iodination of 2'-fluoro-3',5'-di-o-acetyl-2'-deoxyuridine (**3.2**)

Step 2 (figure 3.18) is the addition of an iodine in the C5 position on the nitrogenous base on the uridine. It is common in oligonucleotide modification to halogenate that position so that other functional groups can be substituted for it, allowing for more interesting and different structures not found in nature to be created.
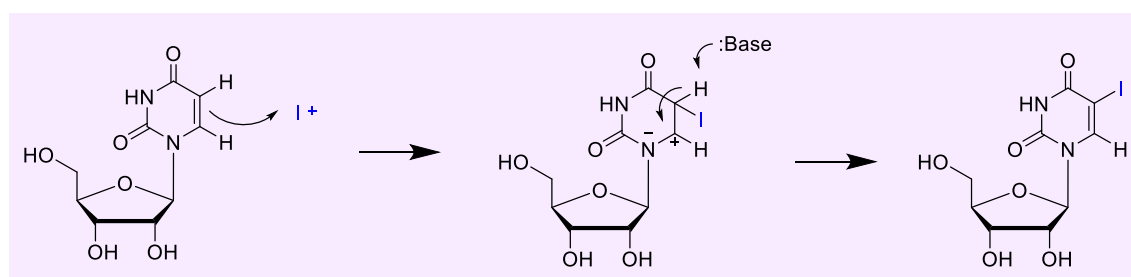


Figure 3.19. Iodination of uridine.[424]

This iodination reaction (figure 3.19) works well. The halogenation reaction occur by attack of an electron-deficient halogen species (formed by in situ oxidation) at C(5) of the uracil ring.[425] It preferentially adds to C(5) as this is the electron-rich position of the substrate.[426] The product is purified to remove the catalyst $(NH_4)_2[Ce(NO_3)_6]$ by a silica gel column (ethyl acetate/methanol 20:1) to produce an orange oil of 3',5'-di-O-acetyl-2'-desoxy-2'-fluoro-5-iodo-uridine (**3.3**). This reaction is highly regioselective meaning no other regioisomers were found

except the desired product.[427] When the reaction is complete there is a peak shift in the $^1$H NMR spectrum. The doublet at 7.71-7.73 ppm shifts to a singlet at 8.19 ppm, this is due to the loss of a hydrogen on the C5 position causing the C6 hydrogen environment to be different. There is a removal of the doublet at 5.69-5.71 ppm from the removal of the hydrogen at the C5 position (figure 3.20). There are extra solvent peaks in the $^1$H NMR spectrum, however it is pure enough to move forward to the next reaction step. The final product was also analysed $^{13}$C NMR (appendix C) and mass spectrometry (appendix D) showing the correct mass peak at 456.98 m/z.
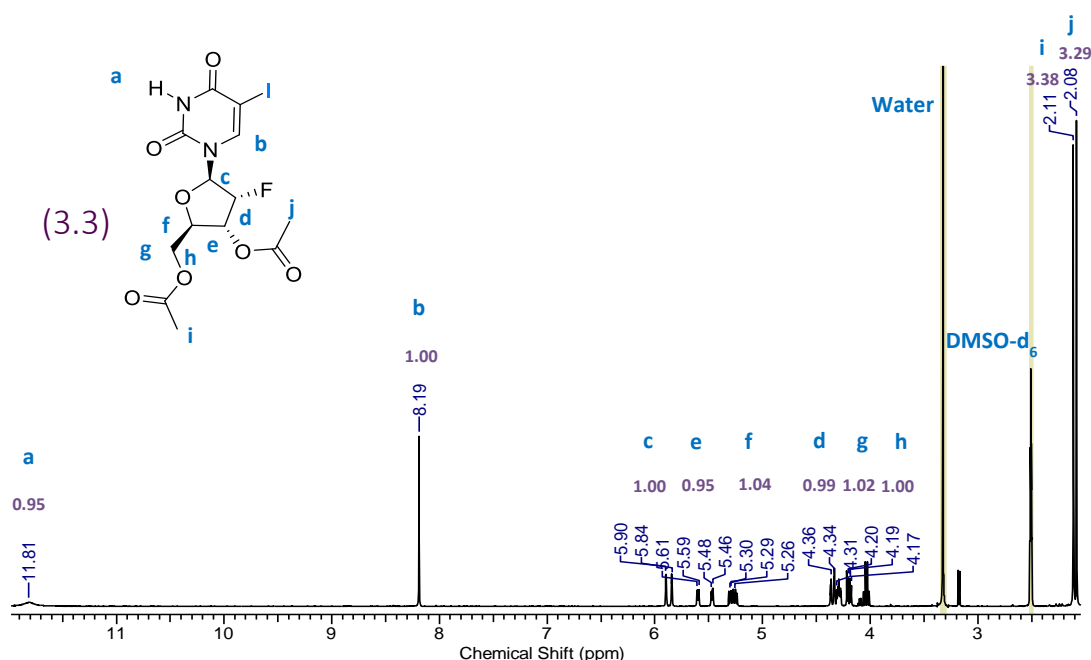


Figure 3.20. $^1$H NMR spectra of 3',5'-di-O-acetyl-2'-desoxy-2'-fluoro-5-iodo-uridine (**3.3**). Extra peak at 3.2 and 4.0 ppm is caused by excess methanol and ethly acetate from the work up of the reaction.

### 3.4.3. Uridine Modification U-Ph

3.4.3.1. Step 3U-Ph: Addition of functional group Ph to 3',5'-di-O-acetyl-2'-desoxy-2'-fluoro-5-iodo-uridine (**3.3**) using Suzuki-Miyaura cross coupling reaction[399]
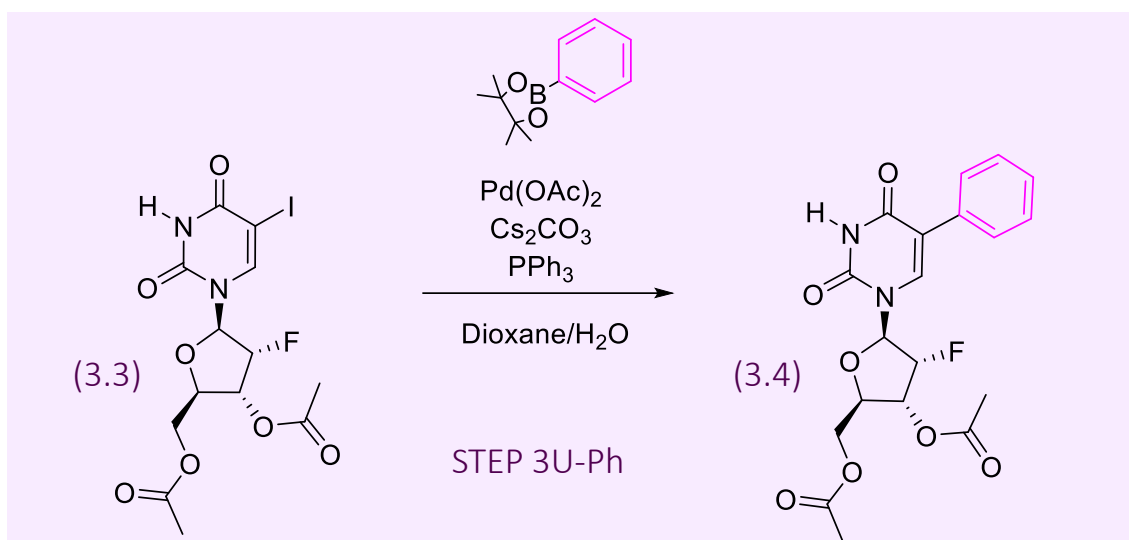


Figure 3.21. Addition of functional group a to 3',5'-di-O-acetyl-2'-desoxy-2'-fluoro-5-iodo-uridine (**3.3**) using Suzuki-Miyaura cross coupling reaction.

Step 3U-Ph shown in figure 3.21 is the addition of functional group Ph. A Suzuki-Miyaura cross coupling reaction using palladium as a catalyst was used to create this modification (figure 3.5). This reaction needed a lot of method and purification development. The palladium catalyst was first filtered out, the product was washed and then a silica gel column (ethyl acetate/methanol = 20:1) was performed. $^1$H NMR analysis of the compound shows a successful coupling by the addition of the aromatic multiplet peaks at 7.32 – 7.74 ppm and a shift of the singlet at 8.19 ppm to 7.94 ppm as the hydrogen environment at position C(6) changes (figure 3.22). There are extra peaks in the spectrum, potentially this compound needs some extra purification steps, however it will be purified in the following reaction steps. The product was analysed by $^{13}$C NMR spectra (Appendix E) and mass spectrometry (Appendix F) showing the correct mass peak plus a Na ion at 429.0 m/z.
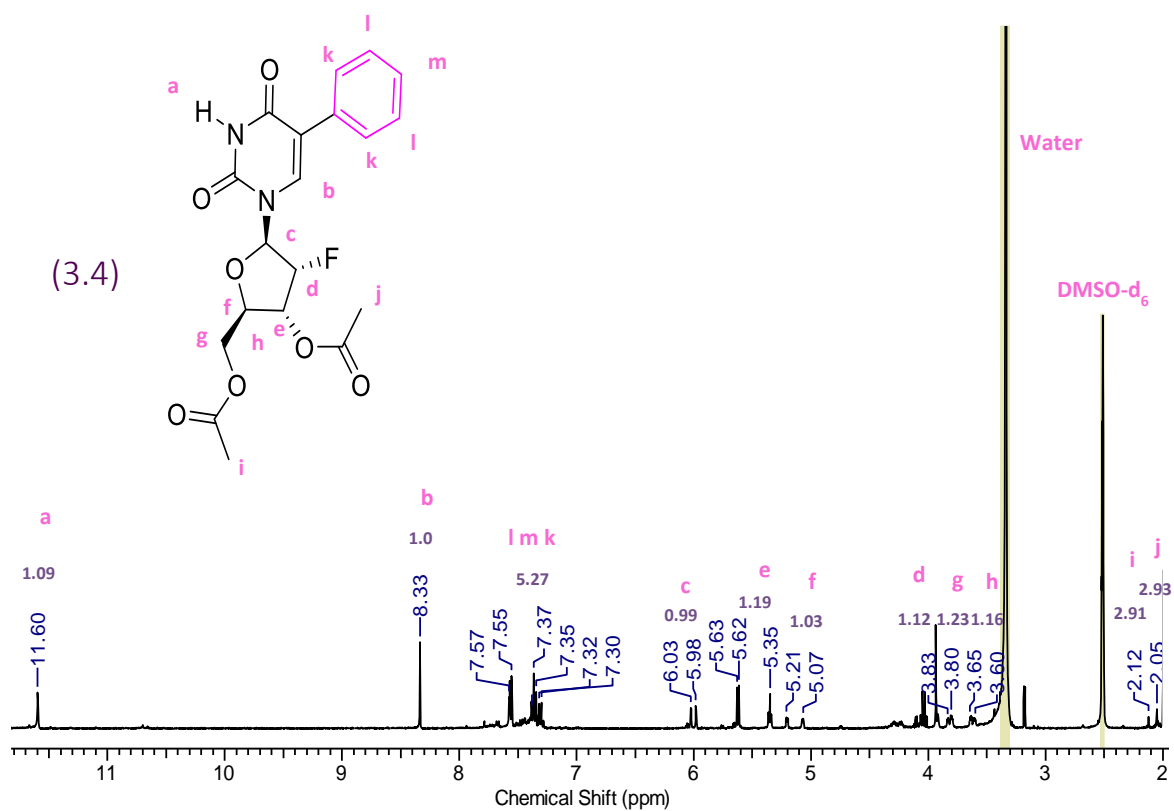
Figure 3.22. $^1$H NMR spectra of 3',5'-di-O-acetyl-2'-desoxy-2'-fluoro-5-phenyl-uridine (**3.4**). Spectrum is not very

clean and there are extra peaks at 3.2 and 4.0 from methanol and ethyl acetate from the work up of this reaction.

## 3.4.4. Uridine Modification U-Vi

3.4.4.1. Step 3U-Vi.1: Addition of functional group Vi to 3',5'-di-O-acetyl-2'-desoxy-2'-fluoro-5-iodo-uridine (**3.3**) using Suzuki-Miyaura cross coupling reaction[399]
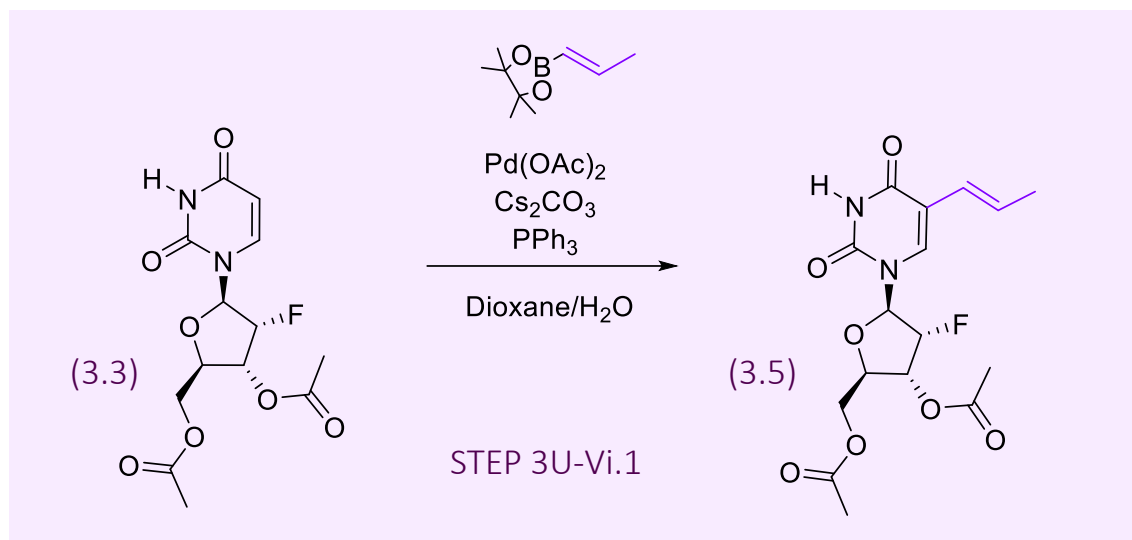


Figure 3.23. Addition of functional group U-Vi to 3',5'-di-O-acetyl-2'-desoxy-2'-fluoro-5-iodo-uridine using Suzuki-Miyaura cross coupling reaction.

Step 3U-Vi (figure 3.23) is the addition of functional group Vi which is the first aliphatic type of modification designed. Functional group Vi is an alkenyl group. A Suzuki-Miyaura cross coupling reaction using palladium as the catalyst was attempted to create this modification. The analysis of this reaction showed that the palladium catalyst was removing the iodine from position C(5) however instead of the new functional group being substituted for it, a hydrogen was bound there. Therefore, the product produced was the same as from step 1, compound **3.2**. The mass spectrometry data show there is not a correct mass peak (figure 3.24). After a lot of experimental attempts and developments to the method functional group Vi was redesigned and a new experimental method planned.

m/z: calcd for $C_{16}H_{19}FN_2O_7$

$([M + H]+)$: 371.12, found 279.1
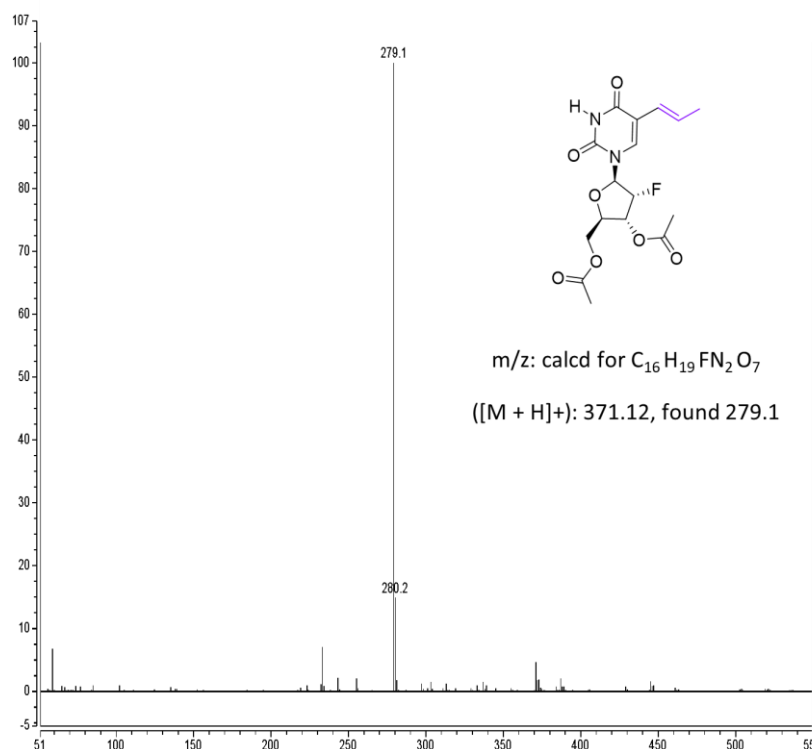
Figure 3.24. Mass spectra of 3',5'-di-O-acetyl-2'-desoxy-2'-fluoro-5-vinyl-uridine (**3.5**).

3.4.4.2. Step 3U-Vi.2: Addition of functional group Vi to 3',5'-di-O-acetyl-2'-desoxy-2'-fluoro-5-iodo-uridine using Stille cross coupling reaction[401]
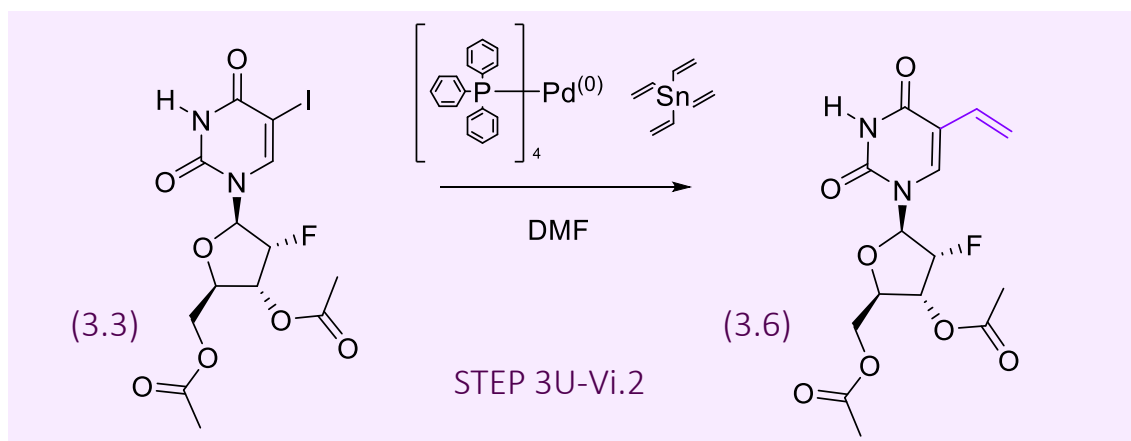


Figure 3.25. Addition of Functional Group U-Vi to 3',5'-di-O-acetyl-2'-desoxy-2'-fluoro-5-iodo-uridine (**3.3**) using Stille Cross Coupling Reaction.

Step 3U-Vi.2 shown in figure 3.25 is the addition of functional group Vi which is the second type of aliphatic modification designed. The revised functional group Vi is an alkenyl group called

159

ethenyl. A Stille cross coupling reaction using tetravinyltin along with tetrakis(triphenylphosphine)palladium(0) as the catalyst was used to achieve this modification (figure 3.8). This Stille coupling reaction worked successfully, without much method development. The purification however needed more enhancement. The palladium and tin both precipitate out of solution and are filtered off. The product was washed and then purified by silica gel column (dichloromethane/methanol = 10:1). $^1$H NMR analysis of the compound shows a successful coupling by the shifting of the signal from C(6) hydrogen from 8.19 ppm down to 7.88 ppm because of the change to the hydrogen environment. There is an addition of a triplet at 5.85-99 ppm and a multiplet at 5.25-23 ppm, from the new hydrogen environments caused by the addition of the ethenyl group on the C5 position (figure 3.26). There is a starting material peak at 8.2 ppm, so more purification steps should be completed on this compound. A mass spectrum shows a peak for the correct mass plus a Na$^+$ ion at 379.00 m/z (appendix H), it was also analysed by $^{13}$C NMR (appendix G).
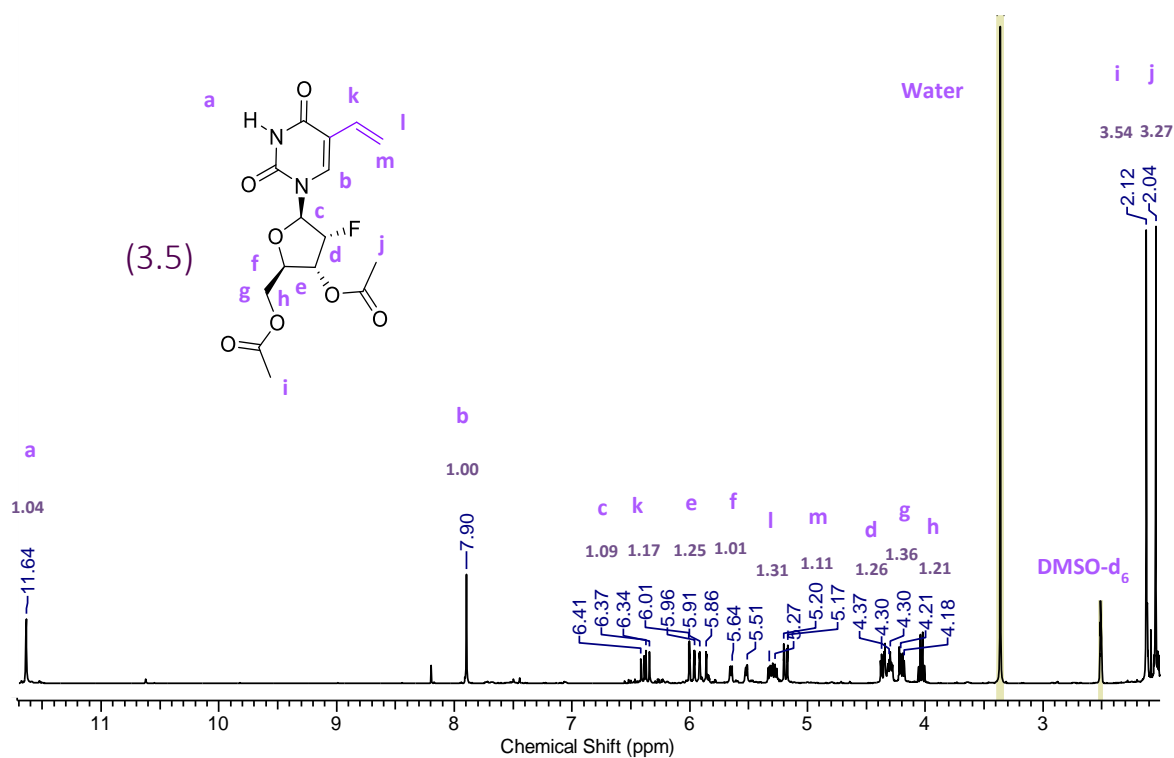
Figure 3.26. $^1$H NMR spectra of 3',5'-di-O-acetyl-2'-desoxy-2'-fluoro-5-vinyl-uridine (**3.6**). There is a starting material peak at 8.2 ppm. More purification steps should be completed on this compound.

## 3.4.5. Uridine Modification U-NH$_2$

3.4.5.1. Step 3U-NH2.1: Addition of functional group **NH$_2$** to 3',5'-di-O-acetyl-2'-desoxy-2'-fluoro-5-iodo-uridine (3.3) using Suzuki-Miyaura Cross Coupling Reaction.
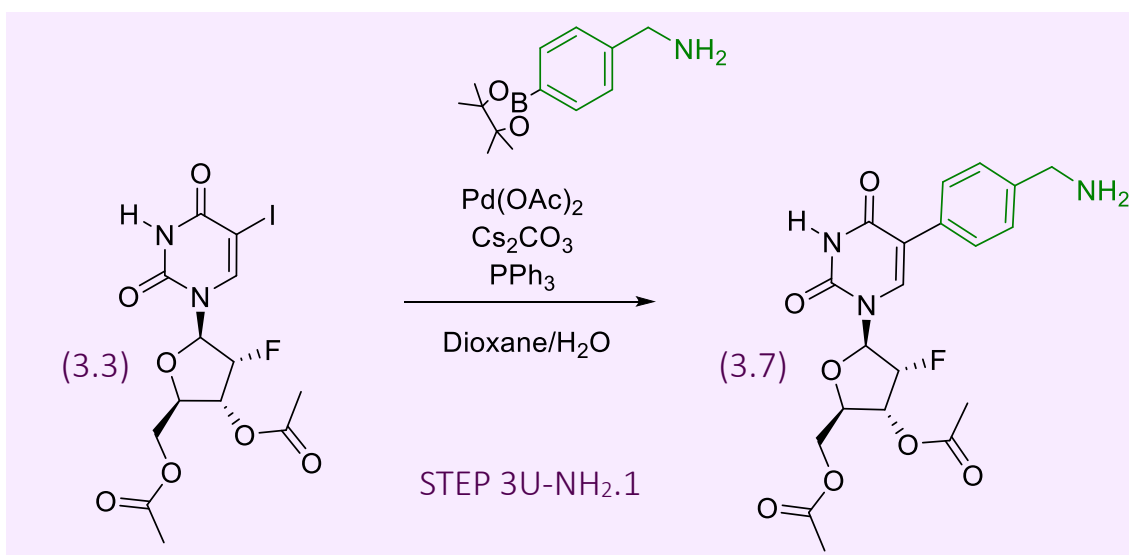


Figure 3.27. Step 3U-NH$_2$.1: Addition of functional group U-NH$_2$ to 3',5'-di-O-acetyl-2'-desoxy-2'-fluoro-5-iodo-uridine using Suzuki-Miyaura cross coupling reaction.

Step 3U-NH$_2$ (figure 3.27) is the addition of functional group **NH$_2$** which is the first amine/ammonium type of modification designed. Functional group **NH$_2$** is an 4-(aminomethyl)phenyl group. A Suzuki-Miyaura cross coupling reaction using palladium as the catalyst was attempted to synthesise this modification. The aim was to add functional group **NH$_2$** on first and then protect it in a second step with a protecting group. The analysis of this reaction showed that the palladium catalyst was removing the iodine from position C(5), however instead of the new functional group being substituted for it, a hydrogen was present. Therefore, the product produced was the same as the one from step 1, compound **3.2**. It was believed that the -NH$_2$ functional group may have been interfering with the coupling, so a protecting group was then added to the 4-(aminomethyl)phenylboronic acid pinacol ester first to try and combat this. The most popular protecting groups used for amines are tert-butoxycarbonyl (Boc), allyloxycarbonyl (Aloc) and fluorenylmethoxycarbonyl (Fmoc).[428] The Fmoc is one of the most used protecting groups and is often chosen as it is compatible with a wide range of linkers.[335] The removal of Fmoc is very efficient, it can be done by treating resin with 20% piperidine in DMF.[429] Fmoc being a base-labile protecting group, works well for nucleic acid synthesis, as the removal of the protecting group can be controlled.[430] Fmoc was chosen to be used in this protection reaction (figure 3.28).
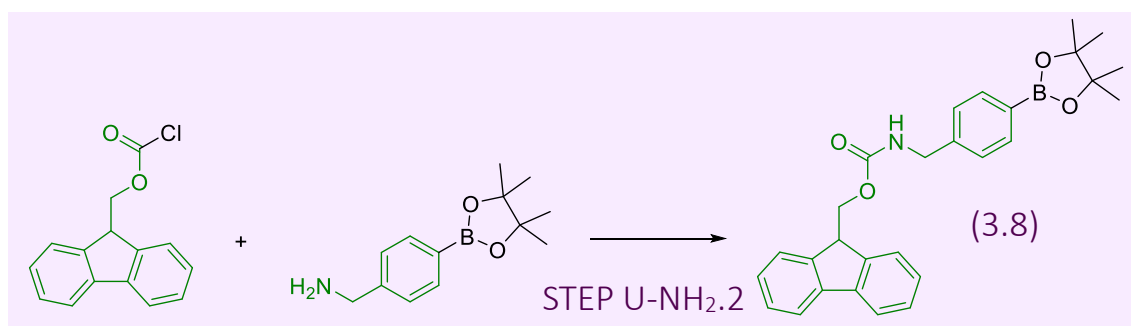


Figure 3.28. Synthesis of Fmoc protected 4-(aminomethyl)phenylboronic acid pinacol ester.

The addition of Fmoc protection group onto 4-(aminomethyl)phenylboronic acid pinacol ester was a successful reaction. $^1$H NMR spectrum (figure 3.29) shows all the corresponding peaks for

this novel product. After the successful synthesis, the experimental procedure was tried again using this novel compound (figure 3.30).
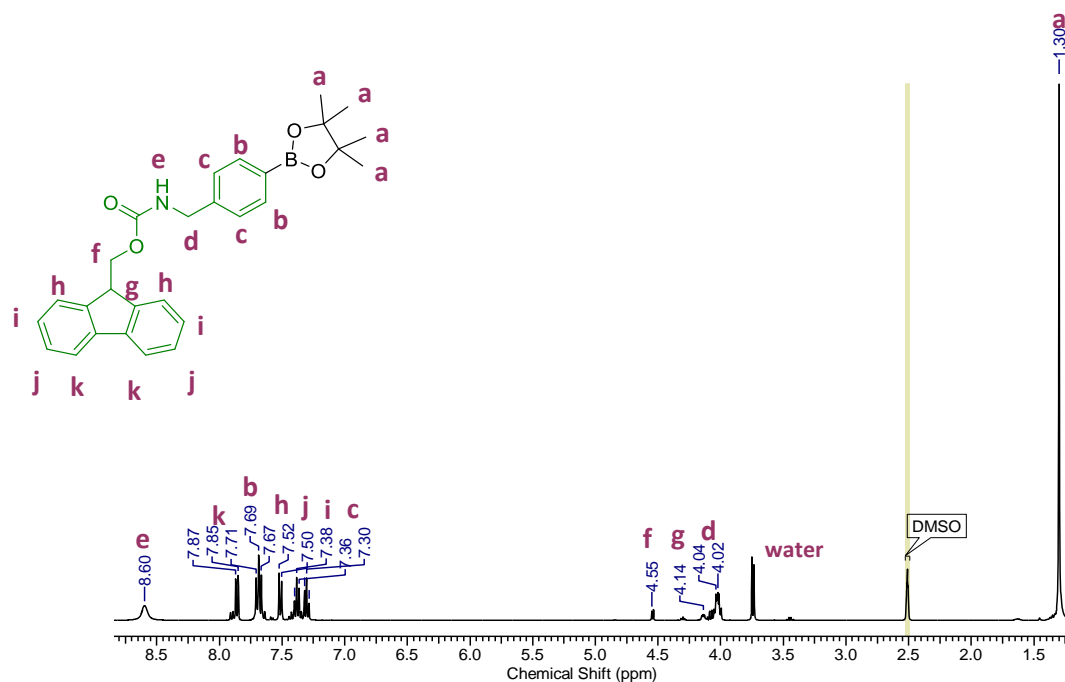


Figure 3.29. $^1$H NMR spectra of FMOC-4-(aminomethyl)phenylboronic acid pinacol ester (**3.8**).
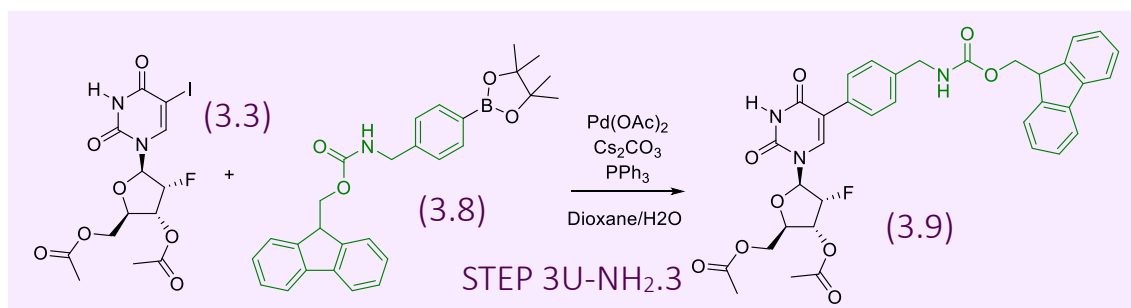


Figure 3.30. Step 3U-NH$_2$.3 Addition of functional group U-NH$_2$ to 3',5'-di-O-acetyl-2'-desoxy-2'-fluoro-5-iodo-uridine (**3.3**) using Suzuki-Miyaura cross coupling reaction.

After a lot of experimental attempts and developments made to the method the coupling was still unsuccessful. The analysis of the final product showed that the palladium catalyst was removing the iodine from position C(5) and substituted it with a hydrogen instead of functional group **NH$_2$**. The mass spectrometry data (figure 3.31) shows that this reaction was unsuccessful. This was believed to be because the protected molecule is now too large causing steric

hindrance, making the coupling of the molecules difficult. After a lot of unsuccessful experimental attempts and method development, functional group **NH₂** was re-designed, and a new route to achieve this was created (figure 3.32).
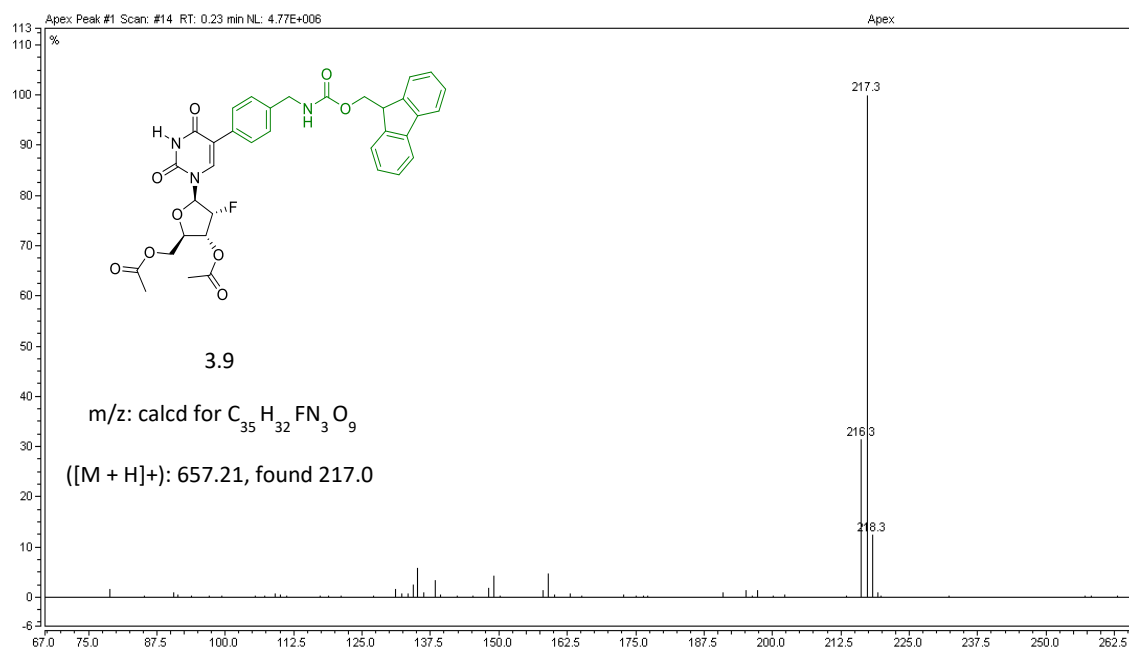


Figure 3.31. Mass spectra of potential product (**3.9**).
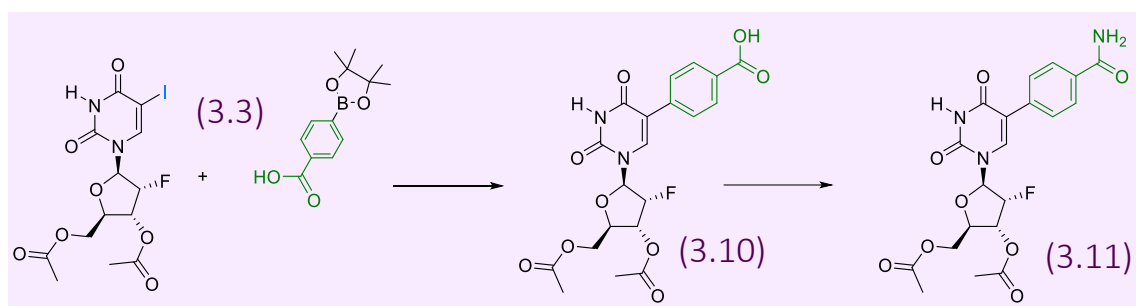


Figure 3.32. New synthesis plan for step 3U-NH₂.4.

The new design of modification **NH₂** was chosen so that the functional group properties were similar to the original design. The aim was to use 4-carboxylphenylboronic acid pinacol ester first, using a similar experimental method as when synthesizing the first modification **Ph**. The carboxylic acid would then be converted into the amide using ammonia solution and dicyclohexylcarbodiimide (DDC) to drive the reaction. If this were successful, then the product would be protected with an Fmoc protecting group.
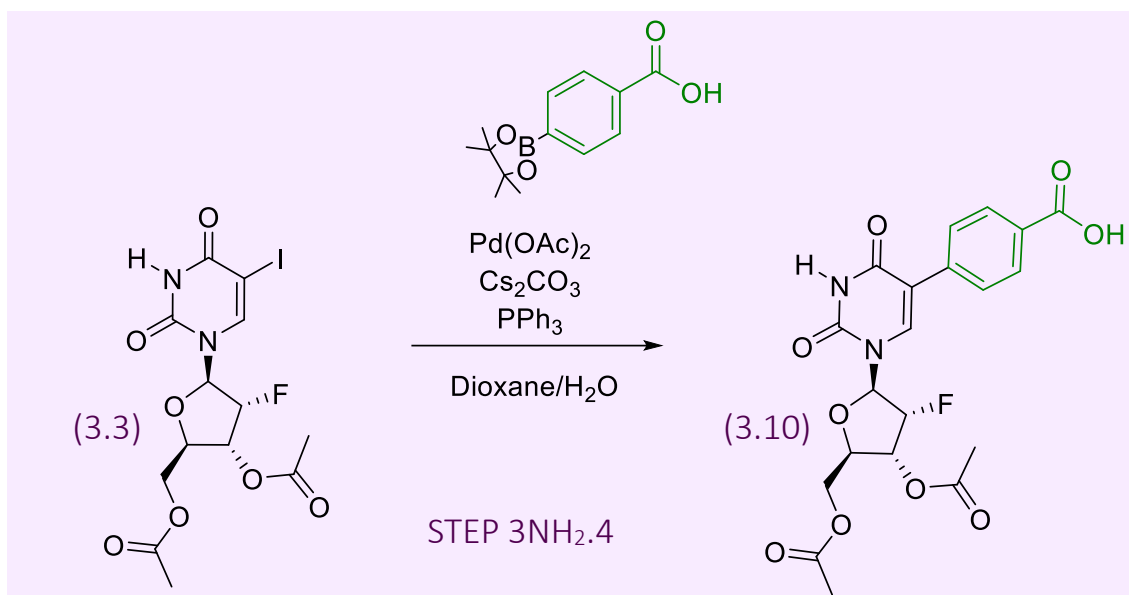
Figure 3.33. Step 3U-NH₂.4 Addition of Functional Group U-NH₂ to 3',5'-di-O-acetyl-2'-desoxy-2'-fluoro-5-iodo-uridine (**3.3**) using Suzuki-Miyaura cross coupling reaction.

Compound **3.10** was not successfully synthesised (figure 3.33). The palladium acetate was removing the iodine and replacing it with a hydrogen atom giving the product compound **3.2**, this can be seen in the COSY NMR (figure 3.34), where the two hydrogen environments at C(5) and C(6) are next to each other. A different palladium catalyst was tried, bis(diphenylphosphino)ferrocenepalladium (II) dichloride[402] without success.
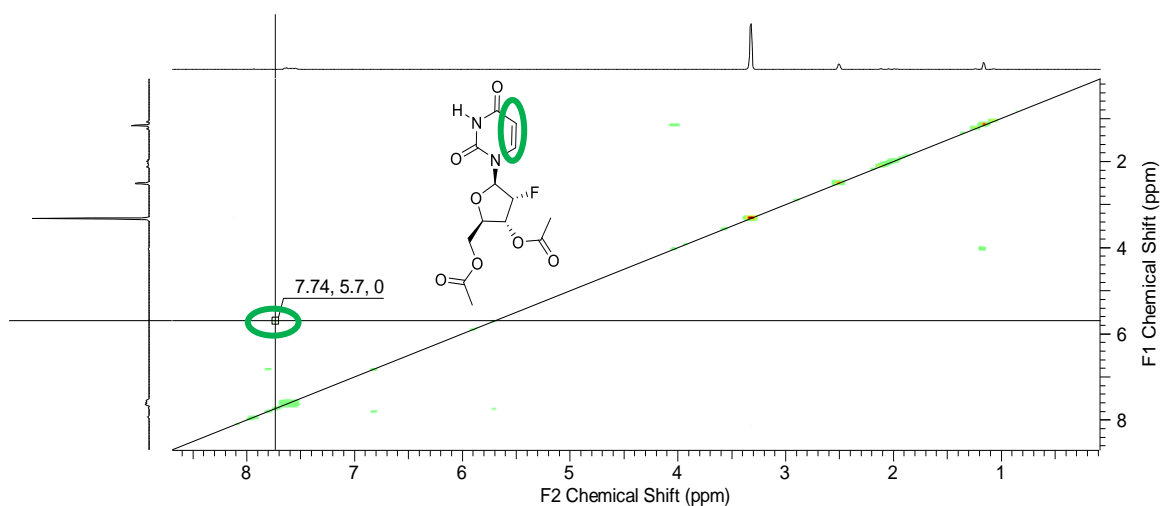


Figure 3.34. COSY NMR spectra of compound **3.10.**

Modification **NH₂** was redesigned, again the new modification was chosen so that the functional group properties were similar. This new design will still create a positive charge like the original amine functional group design. This will be done by a two-step process (figure 3.35). The aim is to nitrate the C5 position instead of an iodination step that was run previously using nitrosonium tetrafluoroborate creating 5-nitro-3',5'-di-*O*-acetyl-2'-fluorouridine. After the nitration reaction, the product would be hydrogenated in the presence of 10% Pd/C, synthesising the 5-amino derivative. This product will then be protected with an Fmoc protecting group.
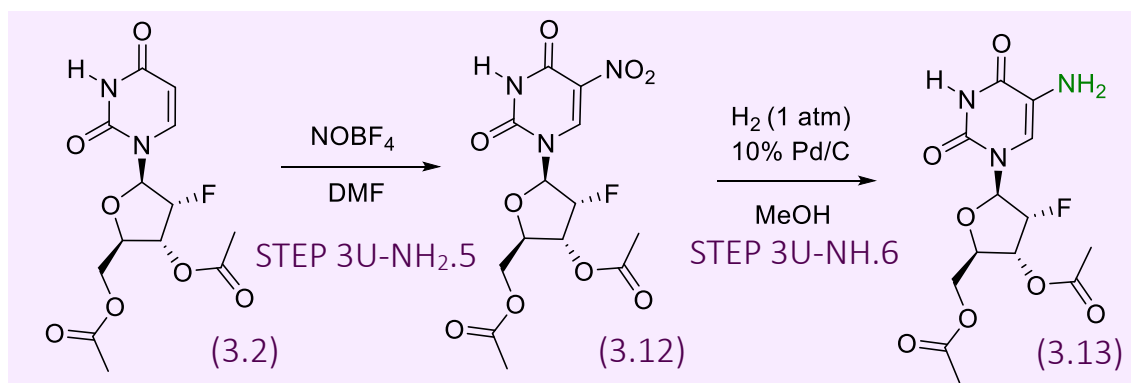


Figure 3.35. Synthesis Plan for 3U-NH₂.5 and 3U-NH₂.6from 3',5'-di-O-acetyl-2'-desoxy-2'-fluoro-5-iodo-uridine (**3.3**).[403,431]

The Step 3U-NH₂.5 was unsuccessful, shown by the mass spectrometry data (figure 3.36) where the mass of 376 m/z is not visible. This reaction could not be followed through to step 3U-NH₂.6.
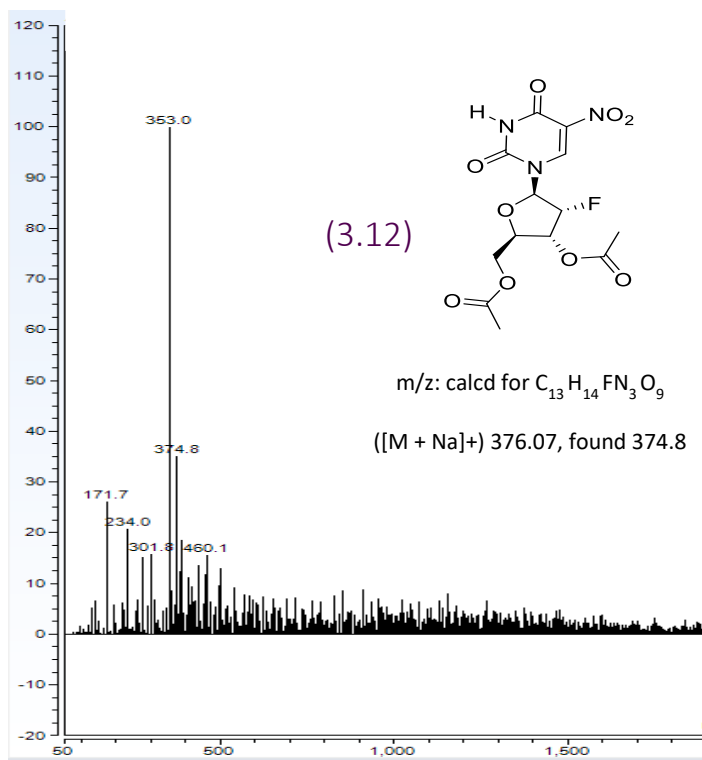
Figure 3.36. Mass spectra of potential product (**3.12**).

After several attempts and method developments, modification **NH₂** was redesigned (figure 3.37).
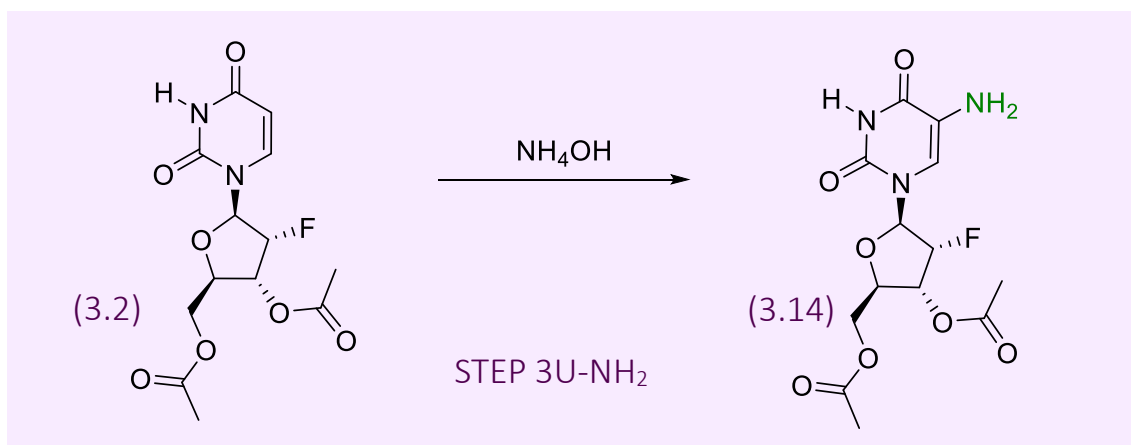


Figure 3.37. Step 3U-NH₂.7 Addition of Functional Group U-NH₂ to 3',5'-di-O-acetyl-2'-desoxy-2'-fluoro-5-iodo-uridine (**3.3**). [404]

The synthesis of compound **3.14** was also unsuccessful; this is shown by the mass spectrometry data (figure 3.38) where the mass of 262.08 m/z is not visible.

m/z: calcd for $C_9 H_{12} FN_3 O_5$

$([M + Na]+)$ 285.05, found 282.0
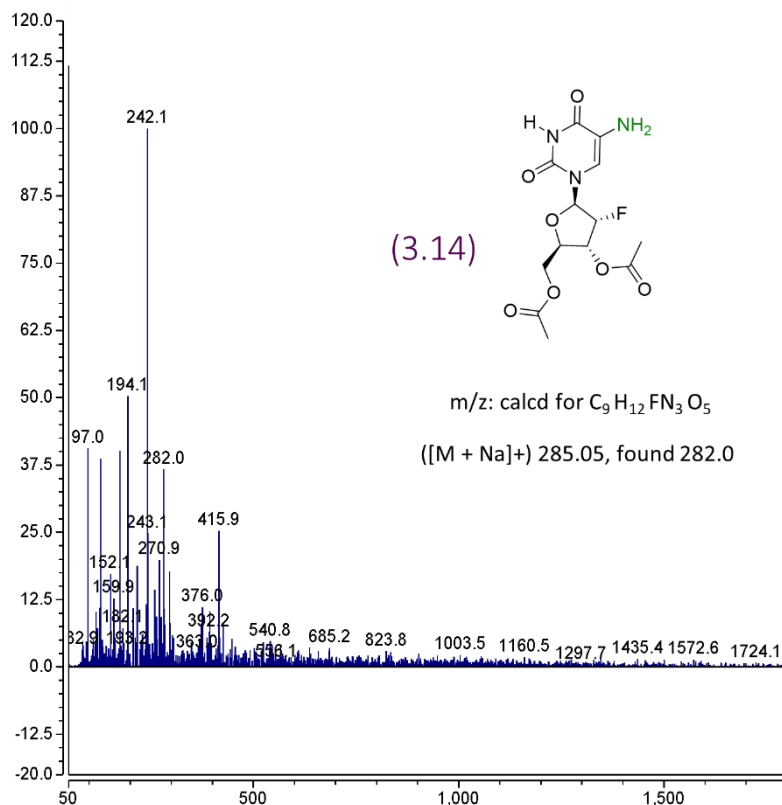
Figure 3.38. Mass spectra of potential product (**3.14**).

## 3.4.6. Uridine Modification U-I

Modification **NH₂** was discontinued. Another modification design was used instead (figure 3.39).
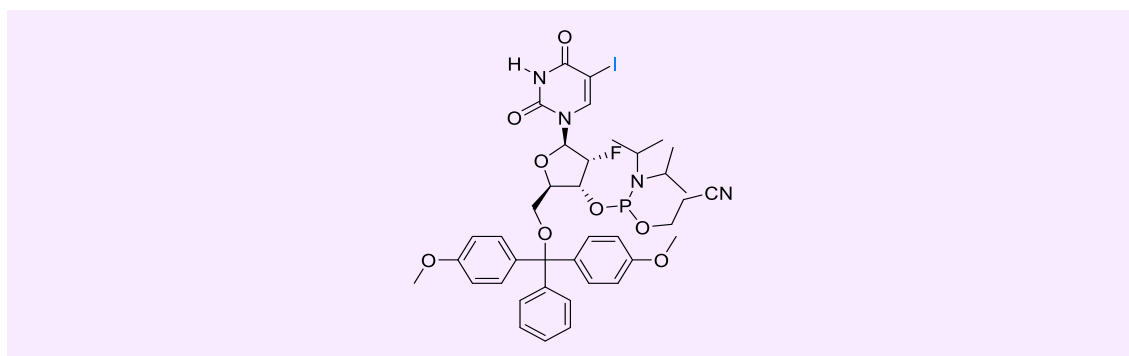


Figure 3.39 Chemical Structure of DMTr -protected 2'-desoxy-2'-fluoro-5-iodo-uridine phosphoramidite (**3.23**)

### 3.4.7. Uridine Modifications, Addition of DMT Protecting group and Phosphoramidite

3.4.7.1. Step 4U-Ph/Vi/I: Removal of acetyl protecting groups from 3',5'-di-O-acetyl-2'-desoxy-2'-fluoro-5-phenyl-uridine (3.4) and 3',5'-di-O-acetyl-2'-desoxy-2'-fluoro-5-vinyl-uridine (3.6) and 3',5'-di-O-acetyl-2'-desoxy-2'-fluoro-5-iodo-uridine (3.3).
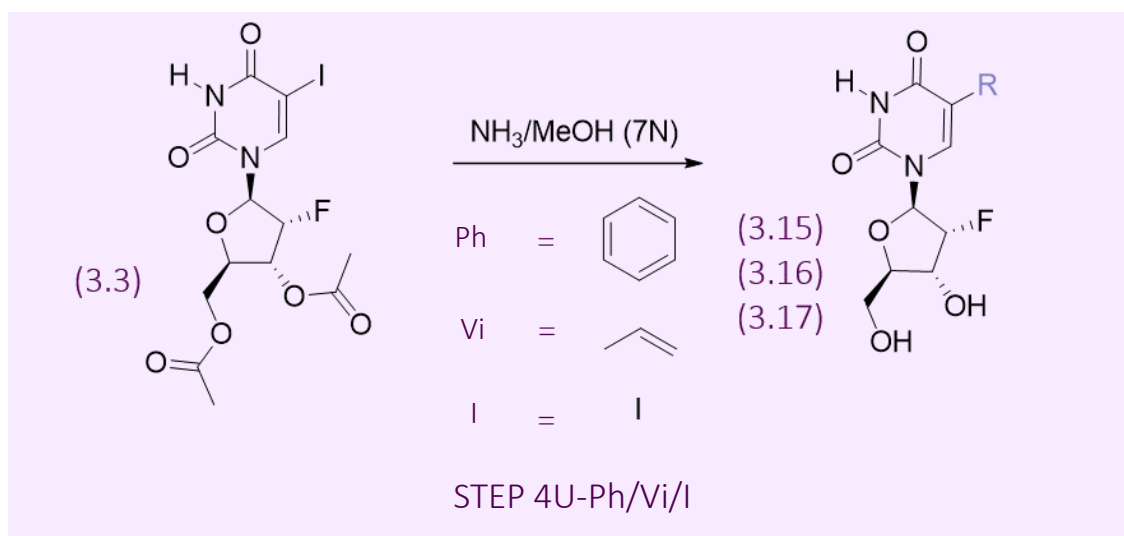


Figure 3.40. Step 4U-Ph/Vi/I: Removal of acetyl protecting groups from 3',5'-di-O-acetyl-2'-desoxy-2'-fluoro-5-phenyl-uridine (3.4) and 3',5'-di-O-acetyl-2'-desoxy-2'-fluoro-5-vinyl-uridine (3.6) and 3',5'-di-O-acetyl-2'-desoxy-2'-fluoro-5-iodo-uridine (3.3).

After the addition of a functional group **Ph**, **Vi** or **I** at C5 position, the acetyl protecting groups were both removed so that a new protecting group could be added to C'5 position and a phosphoramidite could be added to the C'3 position (figure 3.40). An acetyl protecting group is easily removed with NH₃/MeOH (7N) solution at room temperature. This reaction was successful with simple purification needed. $^{1}$H NMR spectra for **3.15** (figure 3.41) shows an addition of single -OH peaks at 3.93 ppm and 5.06 ppm and the loss of 2 large singlet peaks at 2.05 ppm and 2.12 ppm from the acetyl groups. The extra peaks at 4.00 ppm are from ethyl acetate from the work up of the reaction. The extra peakt at 8.50 ppm is from a unknown

contaminant. This $^1$H NMR shows that the reaction was successful along with appendix J showing the correct mass peak of plus a Na$^+$ ion at 345.0 m/z.
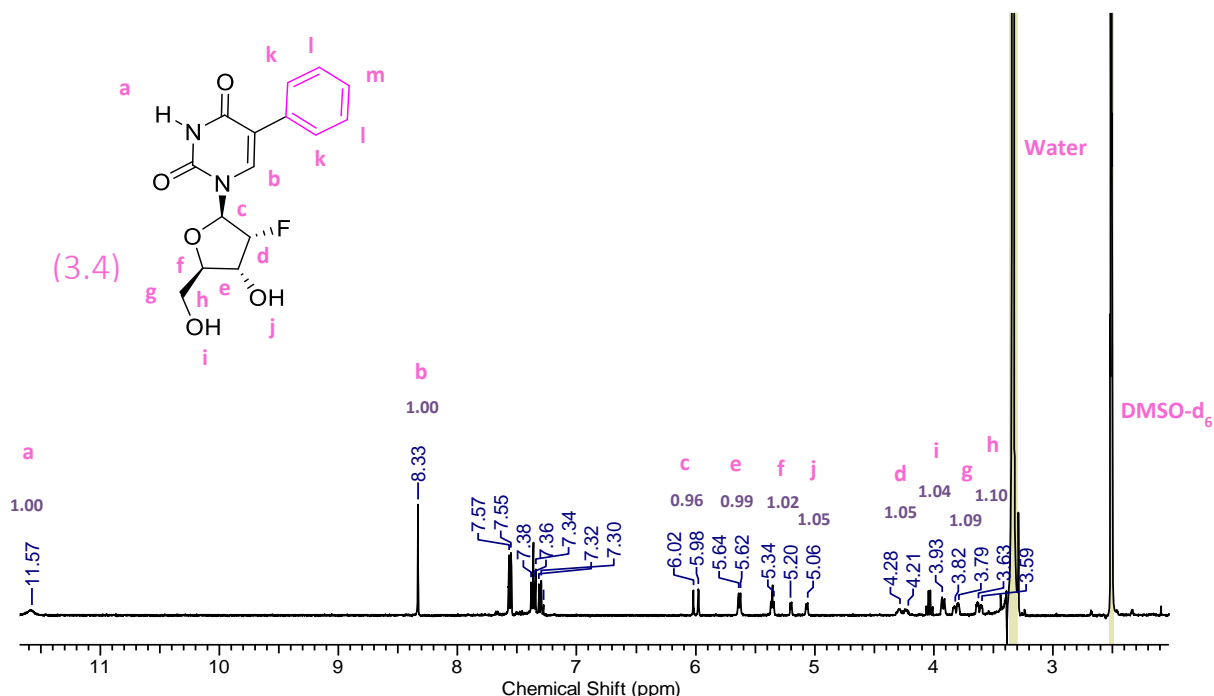


Figure 3.41. $^1$H NMR spectra of 2'-desoxy-2'-fluoro-5-phenyl-uridine (**3.15**). The extra peaks at 4.00 ppm are from ethyl acetate from the work up of the reaction.

$^1$H NMR spectra for **3.16** (figure 3.42) shows an addition of -OH peaks at 3.61 ppm and 4.99 ppm and the loss of 2 large singlet peaks at 2.05 ppm and 2.12 ppm from the acetyl groups. The extra peaks at 4.00 ppm are from ethyl acetate from the work up of the reaction. The extra peakt at 8.50 ppm is from a unknown contaminant. This $^1$H NMR shows that the reaction was successful along with appendix M showing the correct mass peak of plus H$^+$ ion at 272.91 m/z.
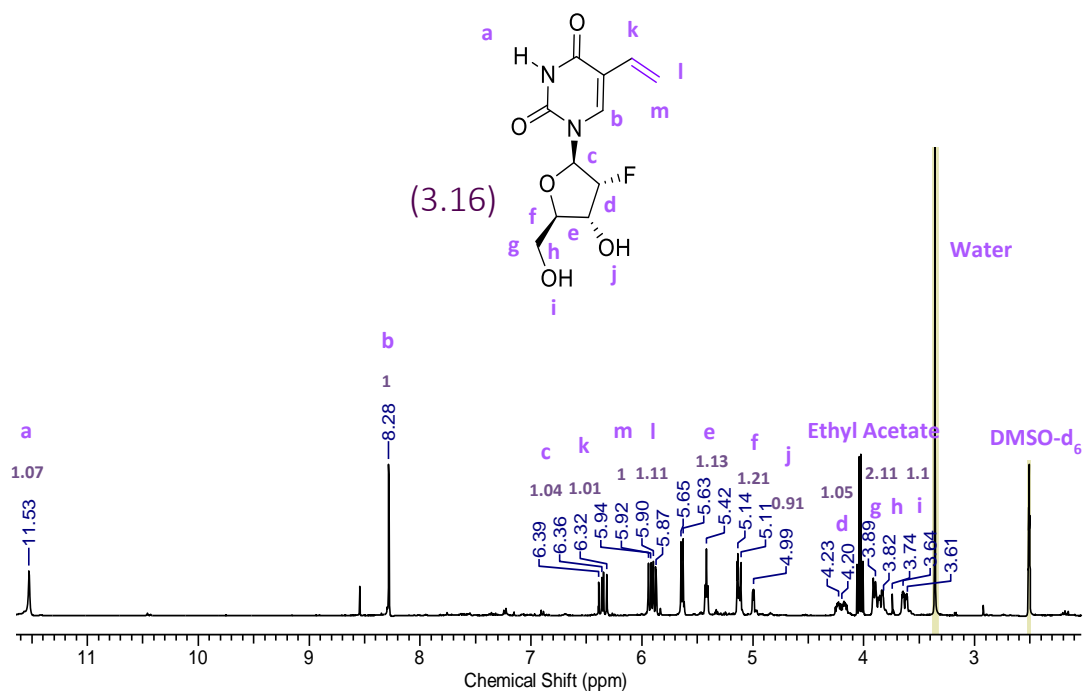
Figure 3.42. ¹H NMR spectra of 2'-desoxy-2'-fluoro-5-vinyl-uridine (**3.16**). The extra peaks at 4.00 ppm are from ethyl acetate from the work up of the reaction. The extra peakt at 8.50 ppm is from a unknown contaminant.

¹H NMR spectra for **3.17** (figure 3.43) shows an addition of single -OH peaks at 3.58 ppm and 5.43 ppm and the loss of 2 large singlet peaks at 2.05 ppm and 2.12 ppm from the acetyl groups. This spectra is clean with no extra solvent peaks. This ¹H NMR shows that the reaction was successful along with appendix O showing the correct mass peak of plus a positive ion at 373.9 m/z.
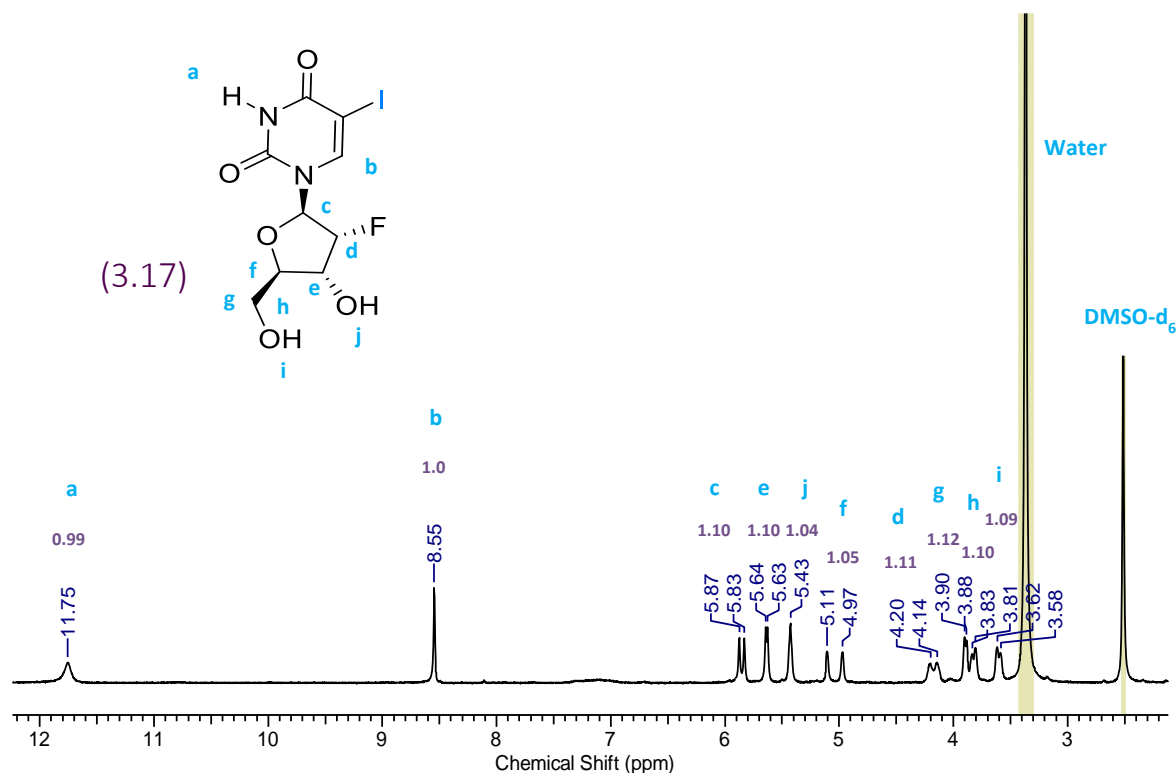
Figure 3.43. ^{1}H NMR spectra of 2'-desoxy-2'-fluoro-5-iodo-uridine (**3.17**).

3.4.7.2. Step 5U-Ph/Vi/I: Addition of 4,4'-dimethoxytrityl (DMT) protecting group to 2'-desoxy-2'-fluoro-5-phenyl-uridine (**3.15**) and 2'-desoxy-2'-fluoro-5-vinyl-uridine (**3.16**)[157] and 2'-desoxy-2'-fluoro-5-iodo-uridine (**3.17**).
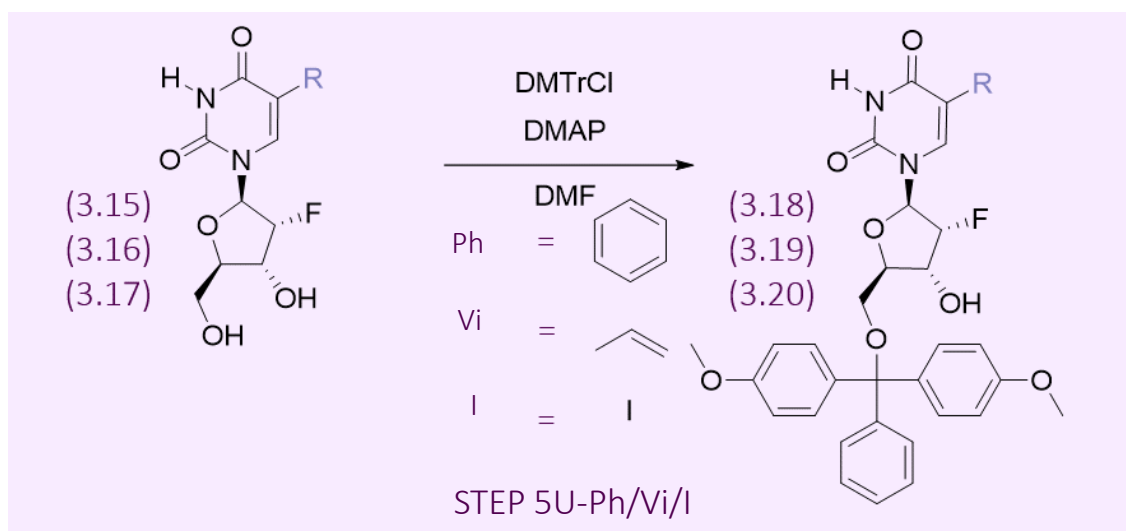


Figure 3.44.  Step 5U-Ph/Vi/I: Addition of 4,4'-dimethoxytrityl (DMT) protecting group to 2'-desoxy-2'-fluoro-5-phenyl-uridine (3.15) and 2'-desoxy-2'-fluoro-5-vinyl-uridine (3.16) and 2'-desoxy-2'-fluoro-5-iodo-uridine (3.17).

After the acetyl protecting groups were removed a 4,4'-dimethoxytrityl protecting group was bound to the 5'-hydroxy group (figure 3.44 and 3.45). The trityl family of protecting groups are acid sensitive hydroxyl protecting groups. These groups are widely used in solid phase organic synthesis for protection and especially in oligonucleotide synthesis. The most popular being dimethoxytrityl (DMT group). 4,4'-dimethyltrityl protecting group must be removed from the nucleoside before the oligonucleotide synthesis can proceed, this is called detritylation, shown below in reaction figure 3[428]. Typically, the de-protection is done with trichloroacetic acid dissolved in dichloromethane. This group is used so that when the samples are put into the DNA synthesiser to synthesise the aptamer chains, the progression can be monitored by the detection of an orange colour that is produced at the visible region 495 nm when the de-protection is successful. The intensity of the orange colour and the absorbance reports the efficiency of the coupling. Most DNA synthesisers have built in hardware that measures the trityl yield for each cycle, so that the efficiency of the synthesis can then be monitored throughout all the cycles[432].
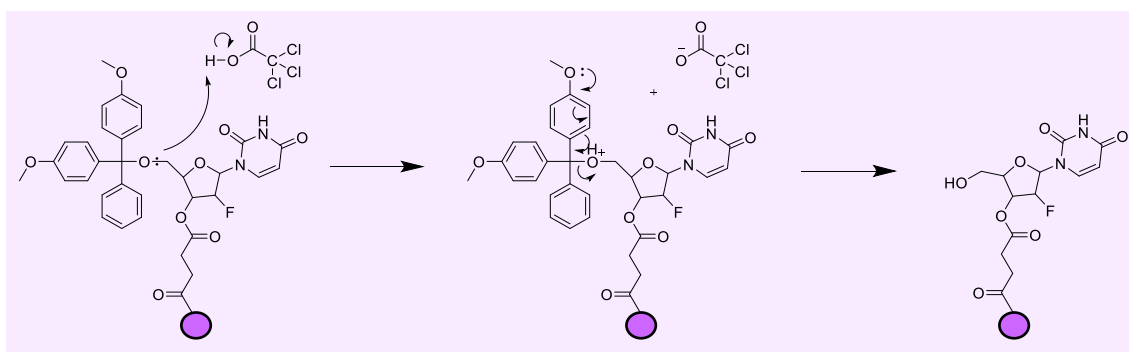


Figure 3.45. First detritylation step of oligonucleotide synthesis.

After a large amount of method development,[433, 434. 435, 436] this reaction worked successfully. The purification process is more complicated, as it must constantly be in basic conditions so as to not cause premature detritylation. Purification has been successful using basic alumina-based columns. $^1$H NMR analysis of the compound **3.18** shows a successful coupling by the addition of aromatic peaks at 6.84-6.86 ppm and 7.19-7.23. And the gain of the large singlet -

CH$_3$ peak at 3.71 ppm (figure 3.46). The extra peaks at 3.20 and 4.00 ppm are from methanol and ethyl acetate from the work up of the reaction. The intergration isn't completely accurat in the aromatic 7.00 ppm region, this maybe be cause by excess DMT group, more purification steps may be required for this step. A mass spectrum shows a peak for the correct mass plus a sodium ion at 647.20 m/z (appendix Q).
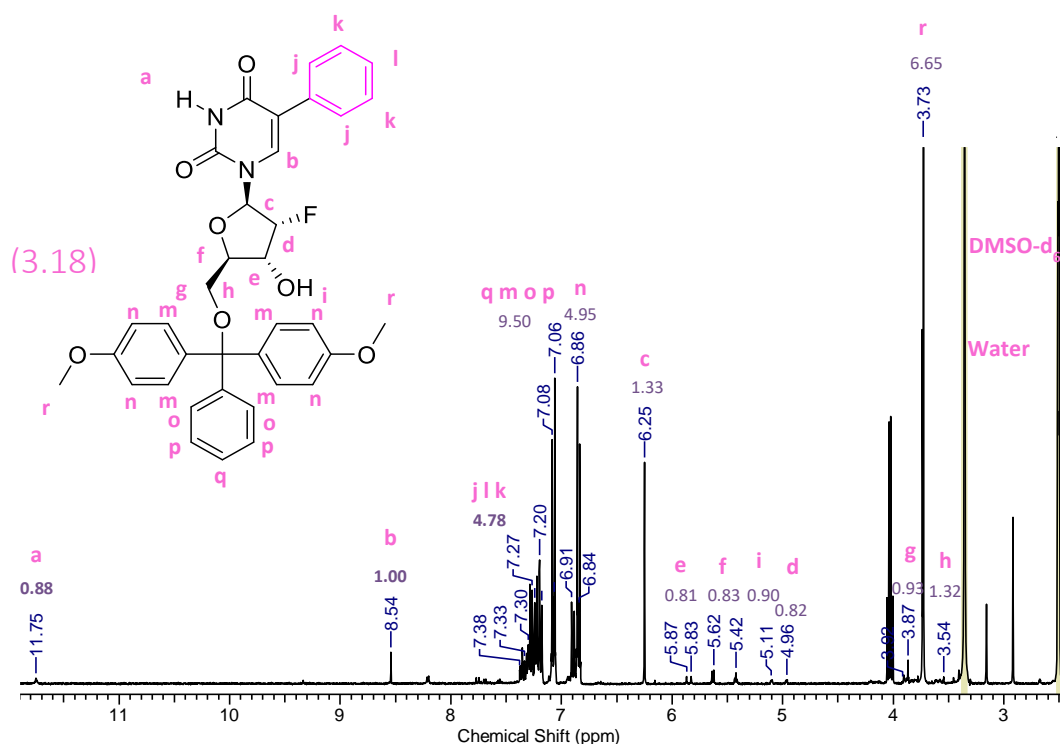


Figure 3.46. ^{1}H NMR spectra of 4'4-DMT tagged 2'-desoxy-2'-fluoro-5-phenyl-uridine (**3.18**). The extra peaks at 3.20 and 4.00 ppm are from methanol and ethyl acetate from the work up of the reaction.

^{1}H NMR analysis of the compound **3.19** shows a successful coupling by the addition of aromatic peaks at 6.86-6.88 ppm and 7.06-7.36. And the gain of the large singlet -CH$_3$ peak at 3.71 ppm (figure 3.47). There are extra peaks on the base line at 7.50 ppm, which may need more purification steps. A mass spectrum shows a peak for the correct mass plus a H ion at 574.00 m/z (appendix U).
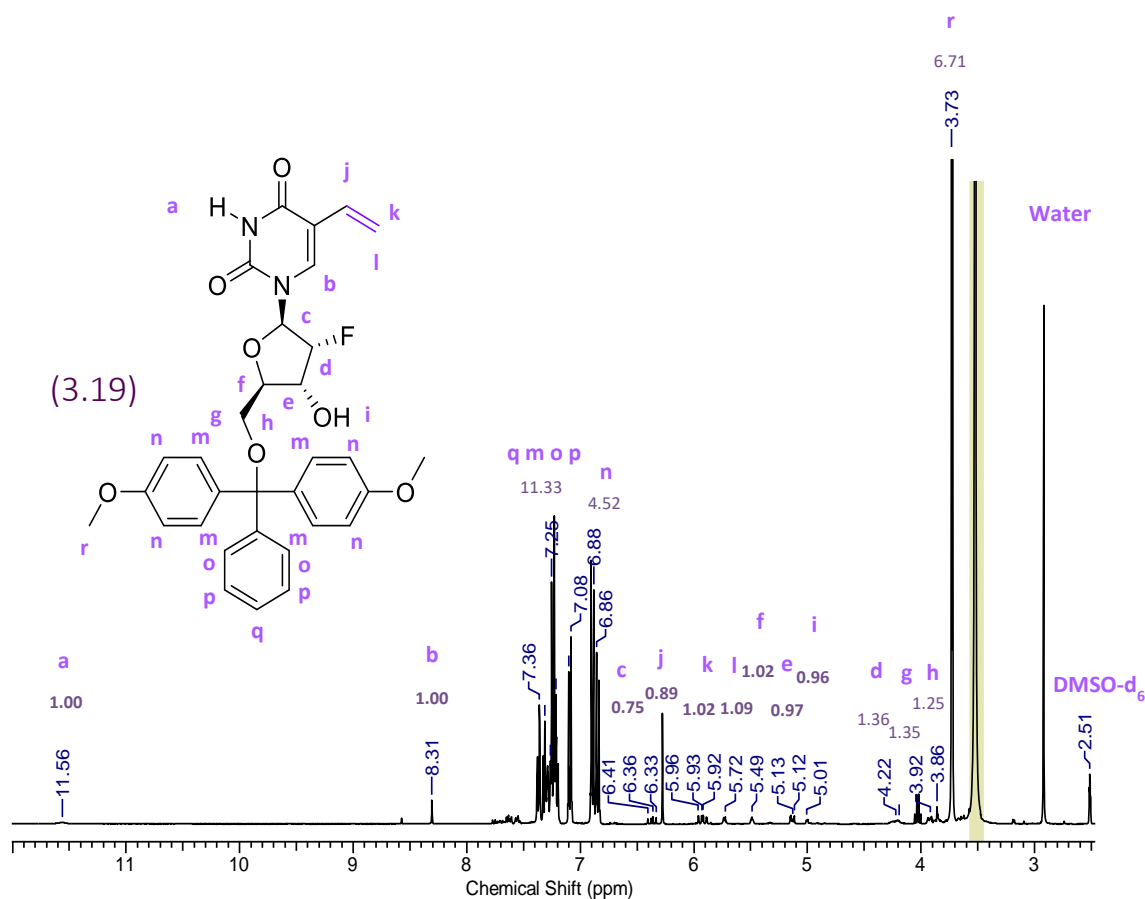
Figure 3.47. ^{1}H NMR spectra of 4'4-DMT tagged 2'-desoxy-2'-fluoro-5-vinyl-uridine (**3.19**). There are extra peaks on the base line at 7.50 ppm, which may need more purification steps.

^{1}H NMR analysis of the compound **3.20** shows a successful coupling by the addition of aromatic peaks at 6.84-6.86 ppm and 7.19-7.37. And the gain of the large singlet -$CH_3$ peak at 3.74 ppm (figure 3.48). The extra peak at 4.00 ppm is from ethyl acetate from the work up of the reaction. The is also an extra peak at 8.05 from starting material that may need more purification steps to remove. A mass spectrum shows a peak for the correct mass plus a $Na^+$ ion at 697.00 m/z (appendix Y).
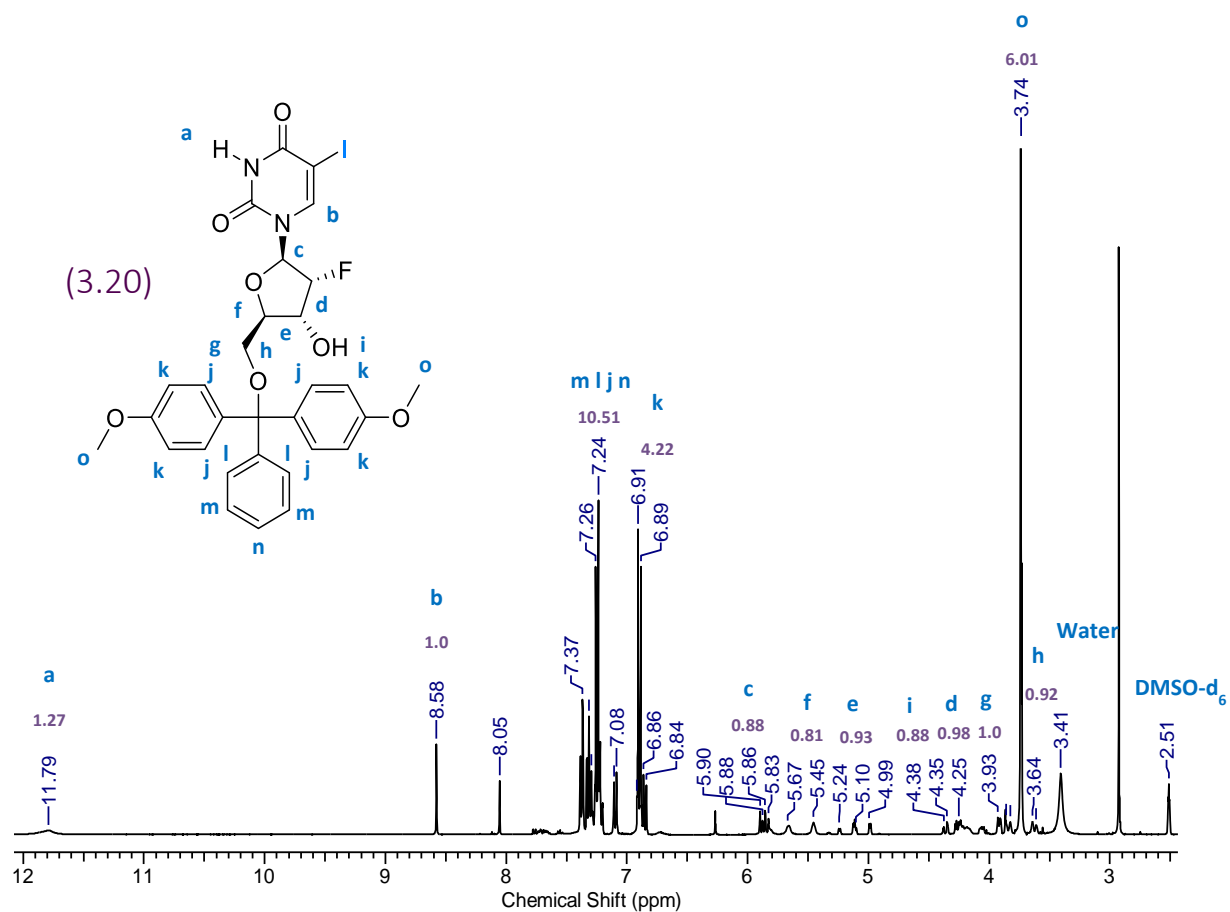
Figure 3.48. $^1$H NMR spectra of 4'4-DMT tagged 2'-desoxy-2'-fluoro-5-iodo-uridine (3.20). The extra peak at 4.00 ppm is from ethyl acetate from the work up of the reaction.

### 3.4.7.3. Step 6U-Ph/Vi/I: Addition of diisopropylchlorophosphoramidite to DMT protected 2'-desoxy-2'-fluoro-5-phenyl-uridine (3.18) and 2'-desoxy-2'-fluoro-5-vinyl-uridine (3.19)[157] and 2'-desoxy-2'-fluoro-5-iodo-uridine (3.20).
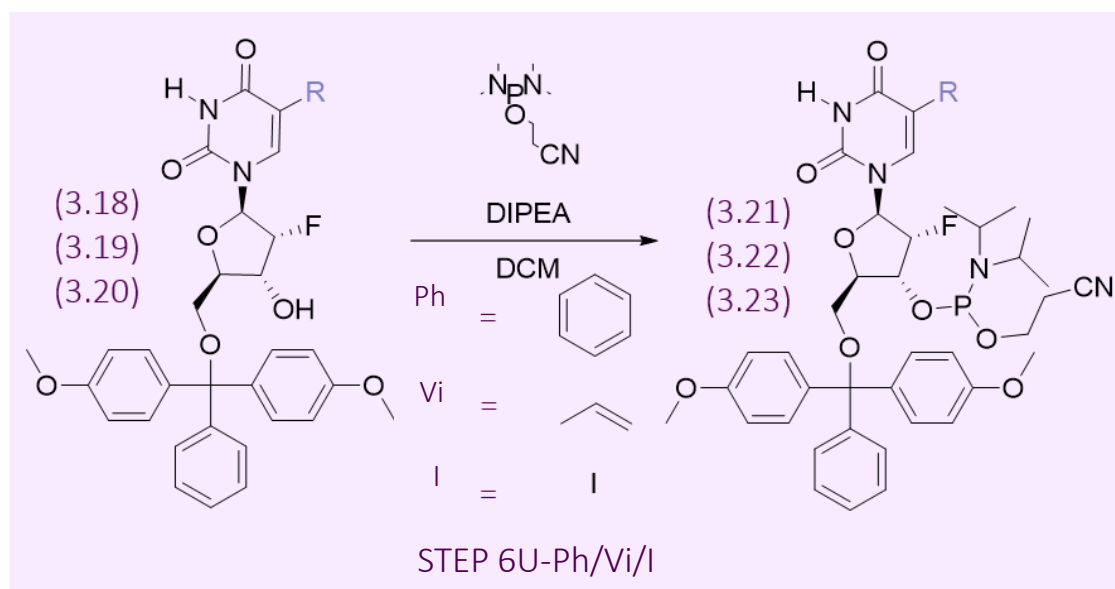


Figure 3.49. Step 6**U-Ph/Vi/I**: Addition of diisopropylchlorophosphoramidite to DMT protected 2'-desoxy-2'-fluoro-5-phenyl-uridine (**3.18**) and 2'-desoxy-2'-fluoro-5-vinyl-uridine (**3.19**) and 2'-desoxy-2'-fluoro-5-iodo-uridine (**3.20**).

The final reaction step before putting the samples on to the synthesiser is to add the diisopropylchlorophosphoramidite, to create the backbone of the RNA aptamer (figure 3.49). In the synthesiser the nucleoside phosphoramidite is added in large excess mixed with an activator, tetrazole. The diisopropylamino group of the nucleoside phosphoramidite is then protonated by the activator which converts it to a good leaving group. This is then quickly displaced by the 5'-hydroxyl group of the nucleoside what is bound to the support on its neighbouring phosphorus atom. This creates a new phosphorus-oxygen bond resulting in a phosphite triester bound to the support (figure 3.50).[437]
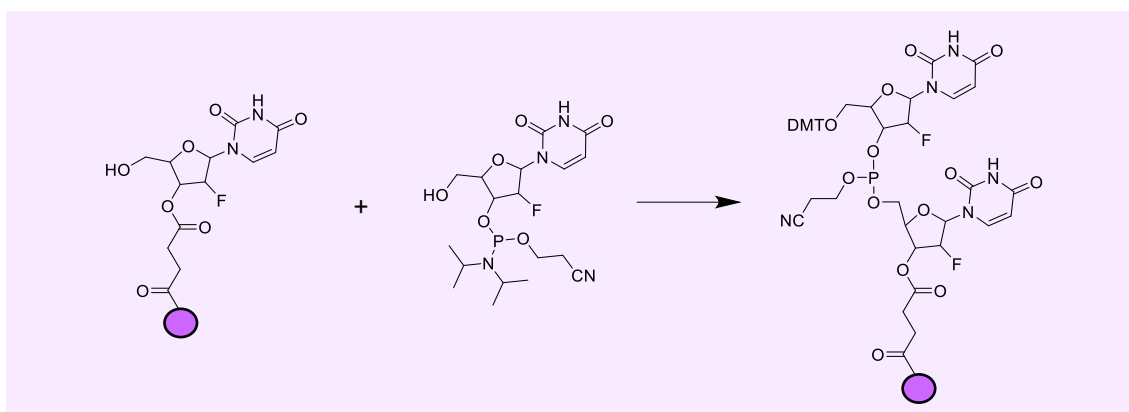
Figure 3.50. Activation and coupling step of oligonucleotide synthesis.

The phosphite-triester P(III) that was created in the coupling step is unstable in acidic conditions so must be converted to the more stable P(V) before the next detritylation step. This conversion is done by iodine oxidation in water and pyridine. This results in the backbone being protected with a 2-cyanoethyl group. This protecting group stops undesirable reactions occurring at the phosphorus during the next cycles (figure 3.51).[348] This step is so important because if the attachment of the diisopropylchlorophosphoramidite is not successful it will prevent the whole synthesis from proceeding.
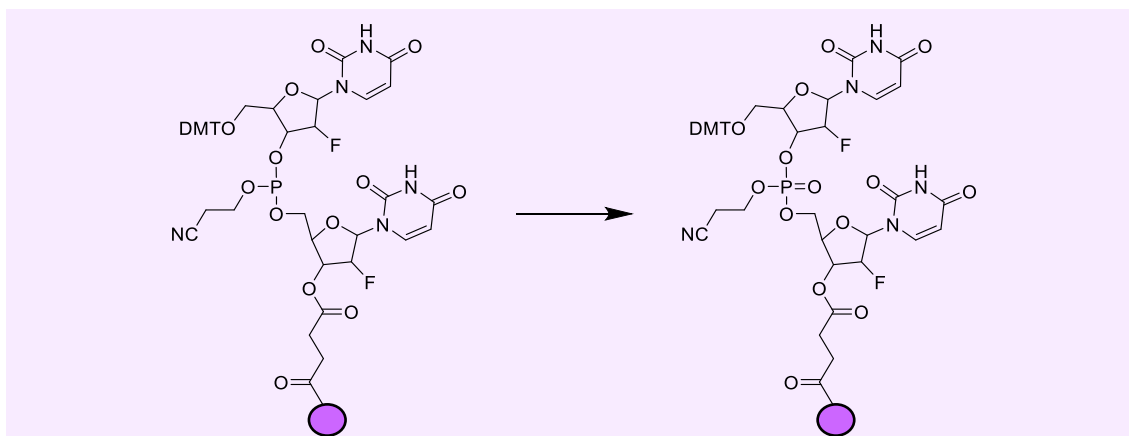


Figure 3.51. Oxidation step of oligonucleotide synthesis.

After some method development this reaction works well and compound **3.18** was put in the synthesiser to check its success. This reaction is air and moisture sensitive, so purification and a long life of the sample has been difficult to achieve. $^{31}$P NMR analysis of the compound **3.21**

shows a successful coupling by the peak at 150.00 ppm (figure 3.52). This final product is a novel

compound, however very similar analogues of this have been made before.
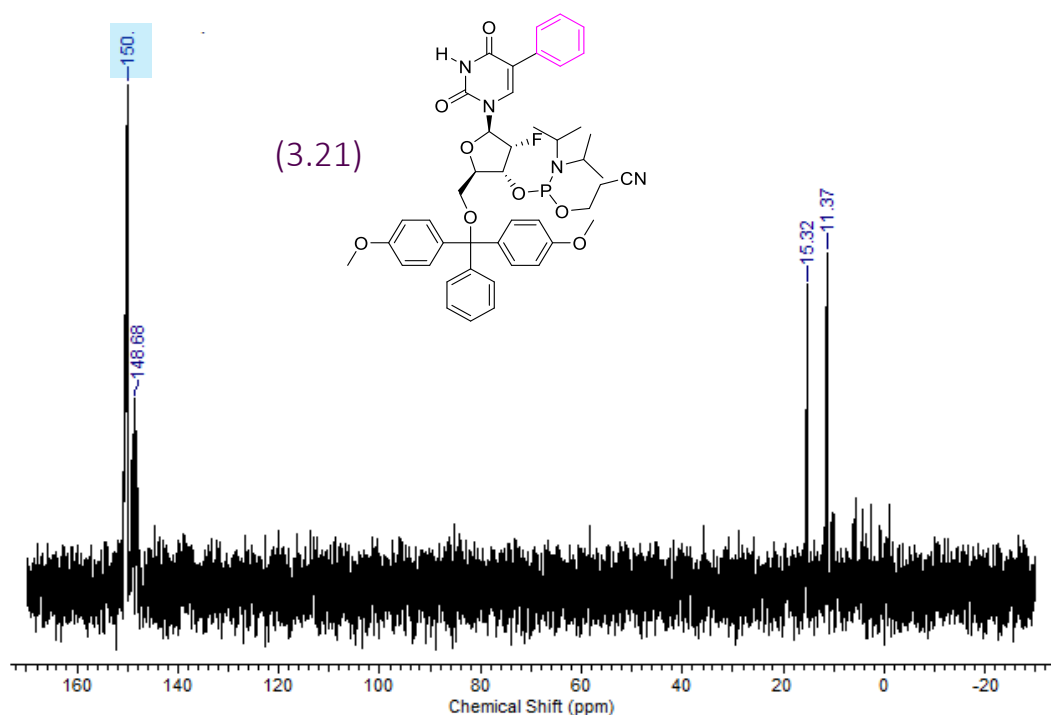


Figure 3.52 ^{31}P NMR spectra of 4'4-DMT tagged 2'-desoxy-2'-fluoro-5-phenyl-uridine phosphoramidite (3.21)

^{31}P NMR analysis of the compound **3.22** shows a successful coupling by the peak at 139.00 ppm

(figure 3.53). The peaks at 11.37-15.32 are oxidized phosphoramidite. The oxidized peaks will

not be purified out of the sample as the purification may cause more oxidation of the sample.

This final product is a novel compound, however very similar analogues of this have been made

before.
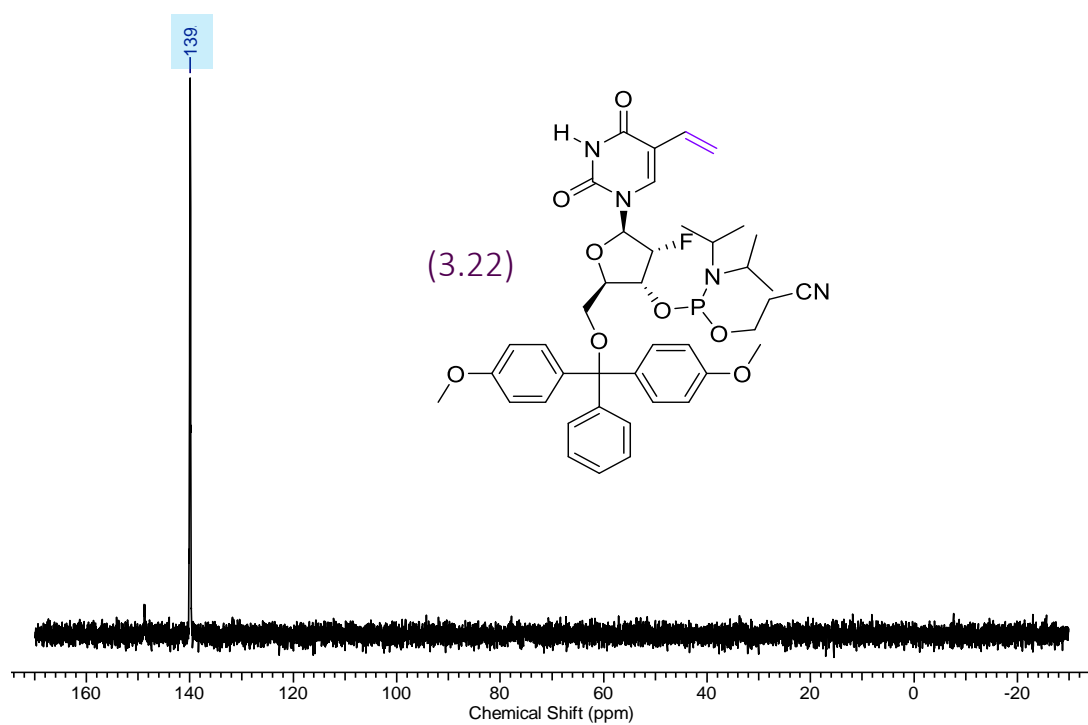
Figure 3.53. ^{31}P NMR spectra of 4'4-DMT tagged 2'-desoxy-2'-fluoro-5-vinyl-uridine phosphoramidite (3.22).

^{31}P NMR analysis of the compound **3.23** shows a successful coupling by the peak at 149.00 ppm (figure 3.54).
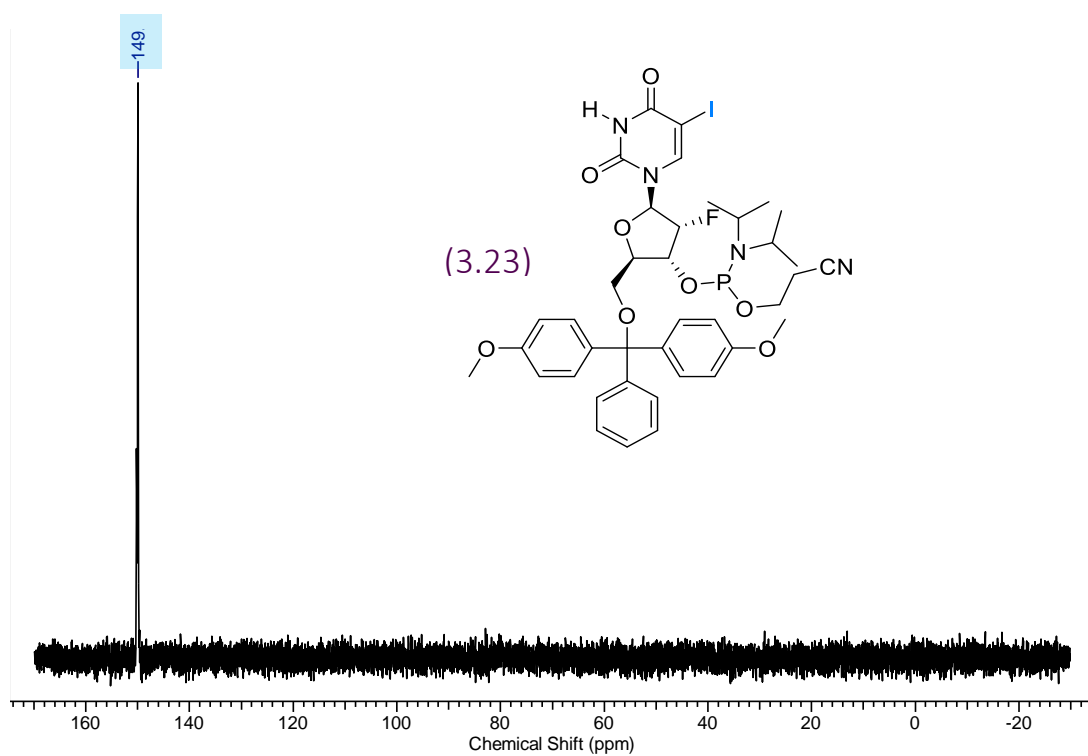


Figure 3.54 ^{31}P NMR spectra of 4'4-DMT tagged 2'-desoxy-2'-fluoro-5-iodo-uridine phosphoramidite (3.23).

## 3.4.8. Quantification of the successful synthesis of a DMT protected 2'-desoxy-2'-fluoro-uridine phosphoramidite.
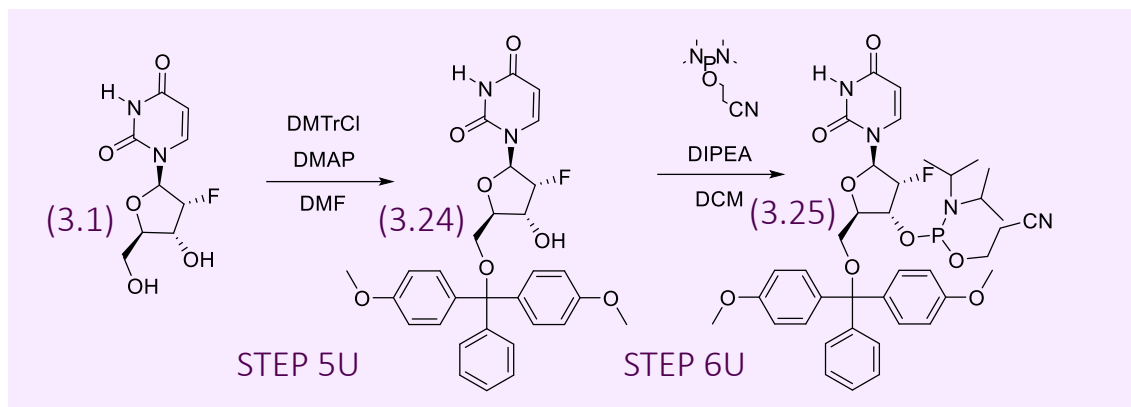


Figure 3.55. Synthesis route of DMT protected uridine phosphoramidite.

To quantify the success of the steps 5U and 6U, unmodified 2'-fluoro-uridine was synthesised to be put on to the Expedite™ DNA synthesiser, to make an aptamer of 25 2'-fluoro-uridines (ARP25U). The sequence being 5'-fUfUfUfUfUfUfUfUfUfUfUfUfUfUfUfUfUfUfUfUfUfUfUfUfU-3'. The full reaction scheme for the synthesis of the oligonucleotide is shown in figure 3.55. This aptamer was synthesised on to a Universal UnyLinker Support as shown in figure 3.56.
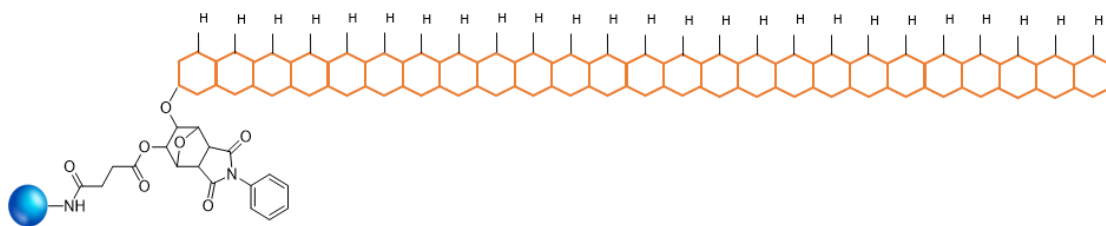


Figure 3.56. ARP25U uridnine aptamer synthesised on Universal UnyLinker Support

The trityl monitor results showed that the 25 2'-fluoro-uridine aptamer synthesis was successful (figure 3.57).
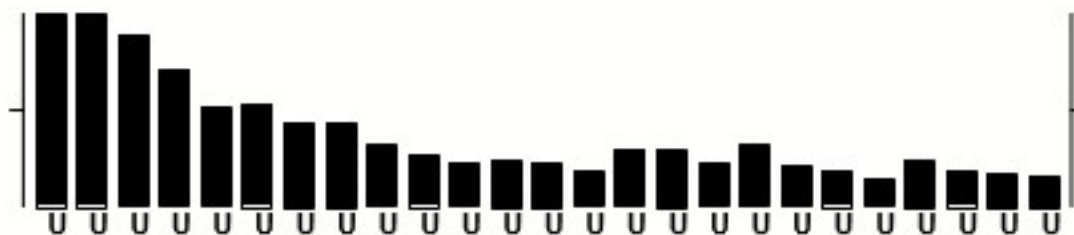
Figure 3.57. ARP25U aptamer trityl monitor progression bar charts.

# 3.5.    Conclusion

In conclusion, phenyl, vinyl and iodo modified uridine phosphoramidite compounds were successfully made and scaled up sufficiently to be used in the synthesis of the MinE07 aptamer library and control aptamers. Successful synthesis of these compounds has been demonstrated by $^1$H NMR, $^{13}$C NMR and mass spectrometry, their purity was good enough to carry forward to the next step. The amino modified uridine phosphoramidite compound was not successfully made. Future work could be done to develop a successful synthesis of this modification.

# 3.6.    Appendix
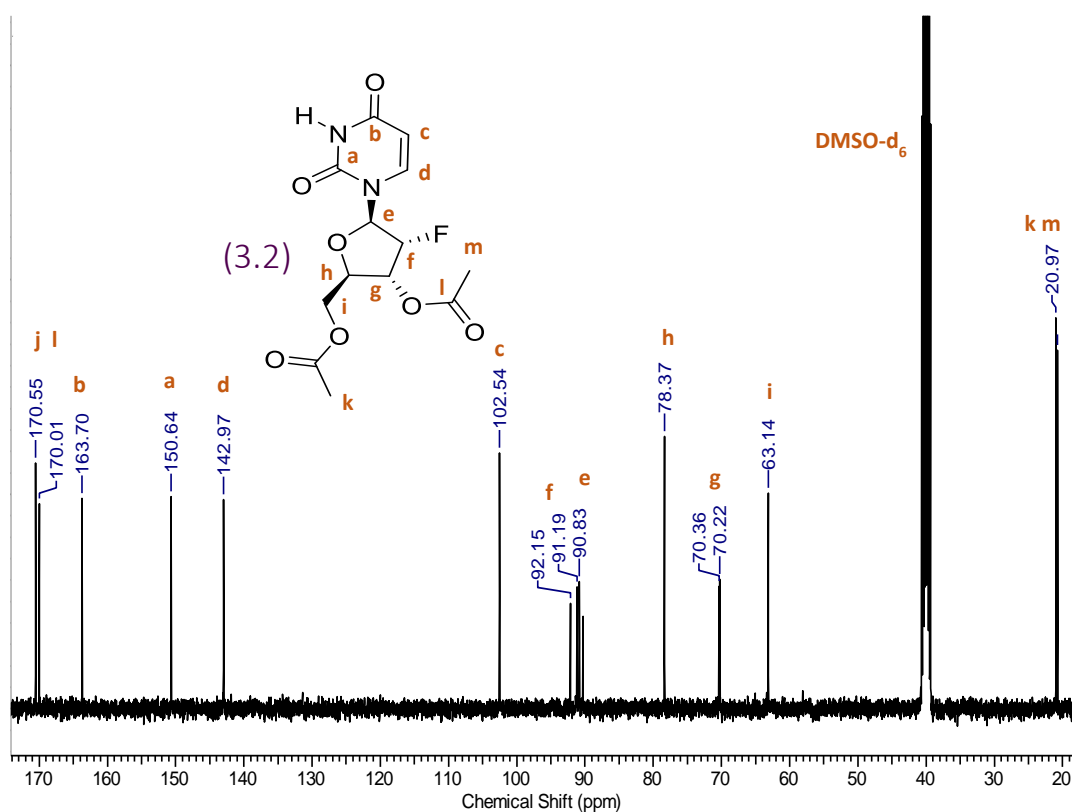


Appendix A: Figure 3.58. $^{13}$C NMR spectra of 2'-fluoro-2'-deoxyuridine (**3.2**).



m/z: calcd for $C_{13} H_{15} F N_2 O_7$

([M + Na]+): 353.09, found 353.04
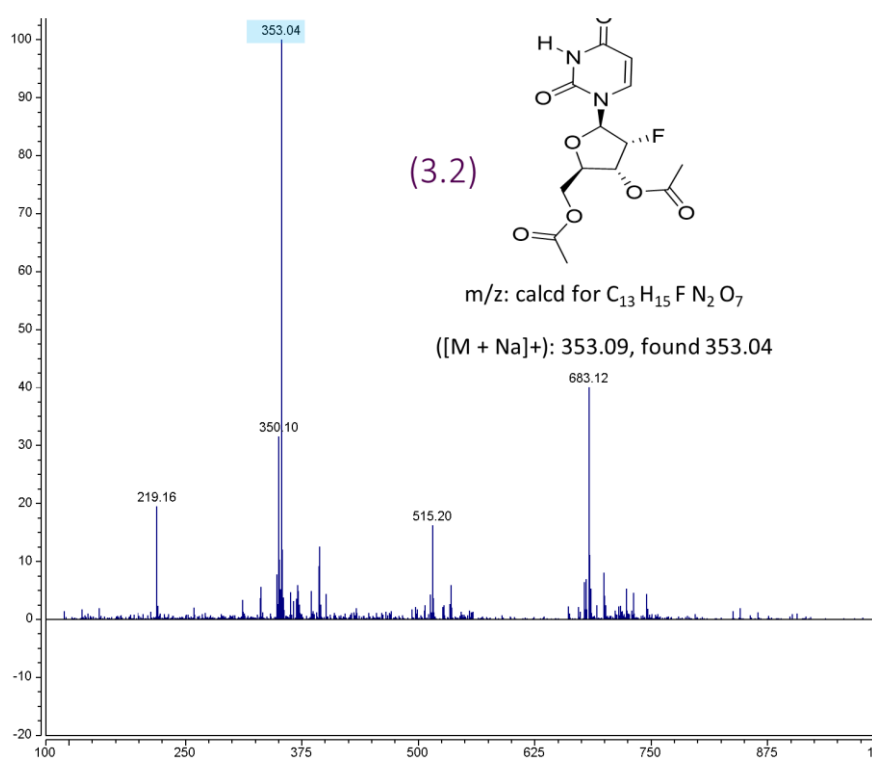
Appendix B: Figure 3.59. Mass spectra of 2'-fluoro-2'-deoxyuridine (**3.2**).
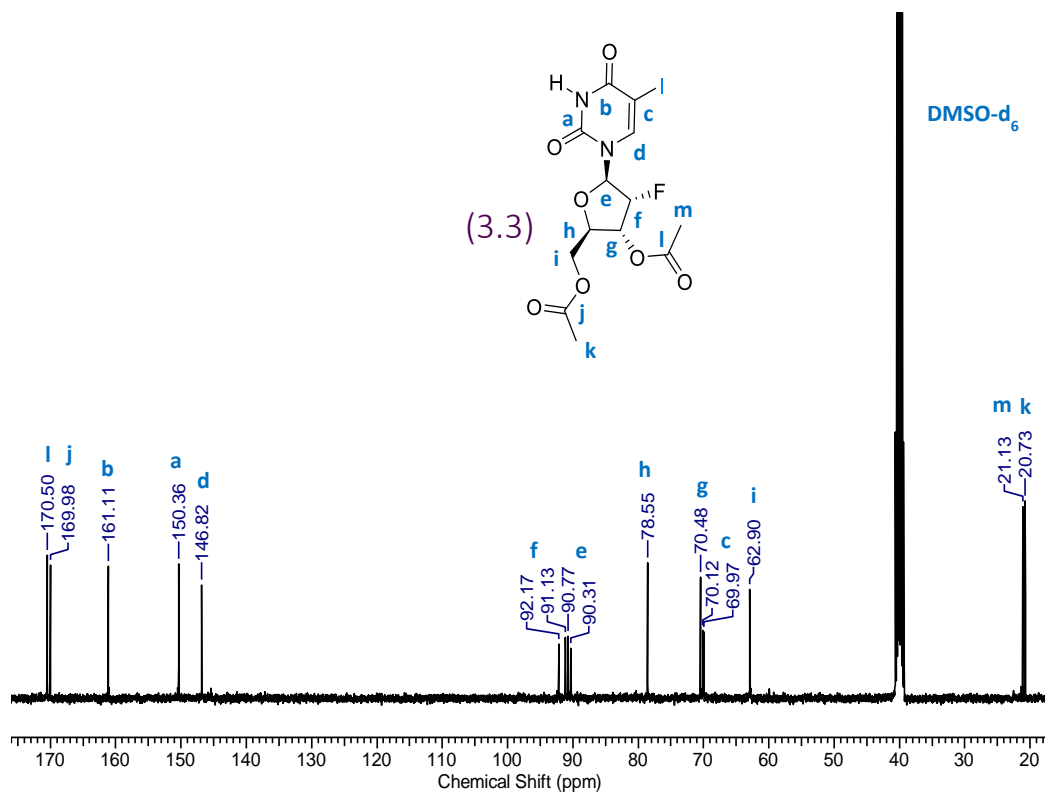
Appendix C Figure 3.60. $^{13}$C NMR Spectra of 3',5'-di-O-acetyl-2'-desoxy-2'-fluoro-5-iodo-uridine (**3.3**).



m/z: calcd for $C_{13}H_{14}FIN_2O_7$
([M + H]+): 456.98, found 456.93
([M + Na]+): 479.98, found 478.92
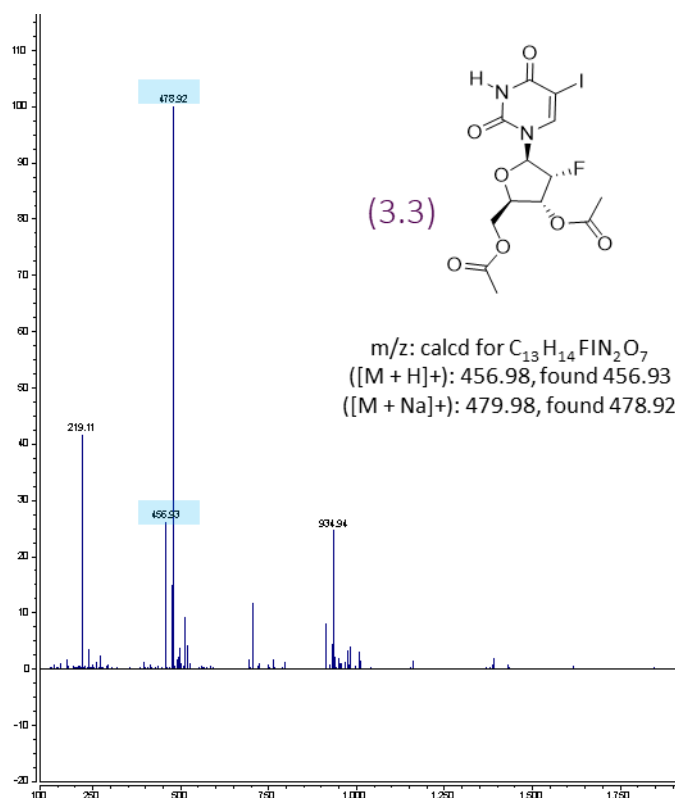
Appendix D: Figure 3.61. Mass spectra of 3',5'-di-O-acetyl-2'-desoxy-2'-fluoro-5-iodo-uridine (**3.3**).
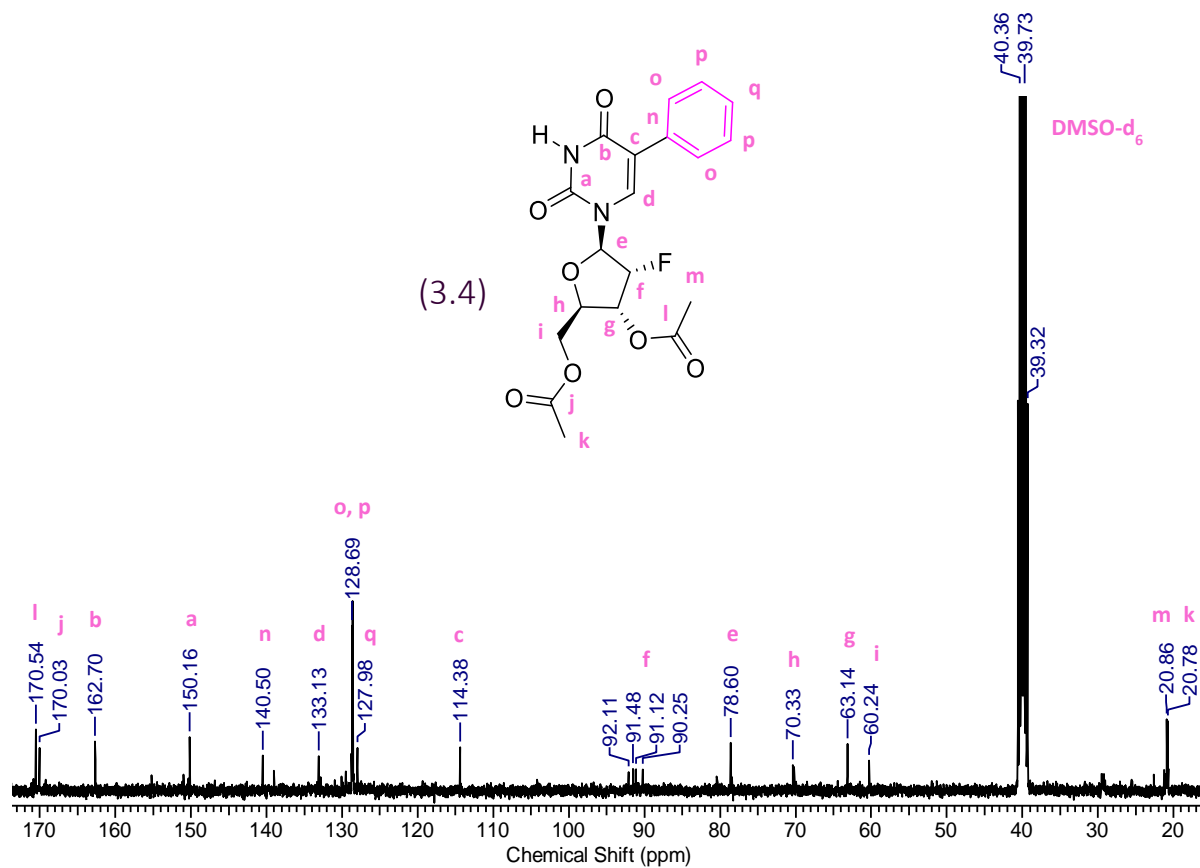
Appendix E: Figure 3.62. 13C NMR Spectra of 3',5'-di-O-acetyl-2'-desoxy-2'-fluoro-5-phenyl-uridine (3.4).



m/z: calcd for $C_{19}H_{19}FN_2O_7$

([M + Na]+): 429.12, found 429.02
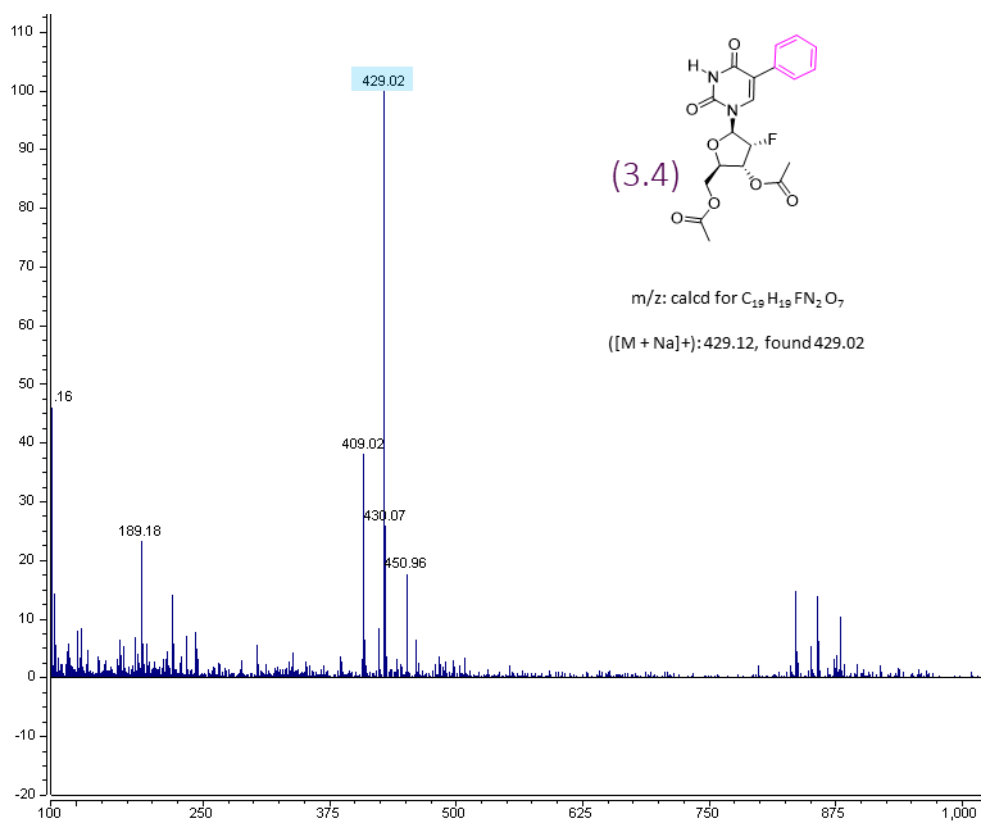
Appendix F: Figure 3.63. Mass spectra of 3',5'-di-O-acetyl-2'-desoxy-2'-fluoro-5-phenyl-uridine (**3.4**).

m/z: calcd for $C_{16}H_{19}FN_2O_7$

([M + H]+): 371.12, found 279.1
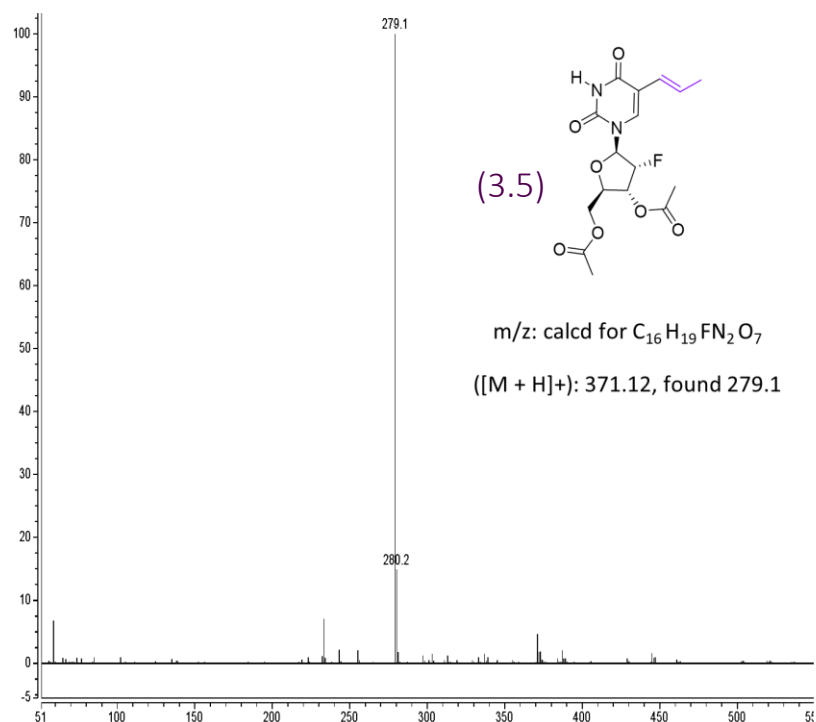
Appendix G: Figure 3.64. Unsuccessful Mass spectra of compound **3.5.**
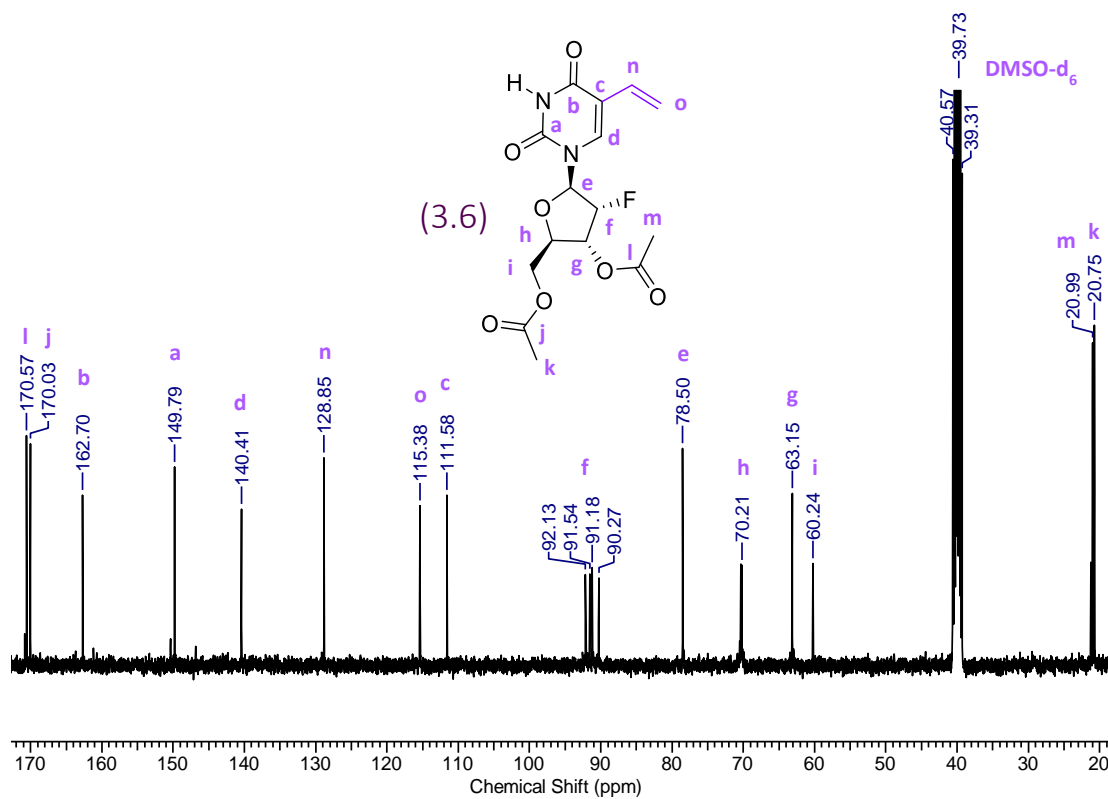


Appendix H: Figure 3.65. $^{13}$C NMR spectra of 3',5'-di-O-acetyl-2'-desoxy-2'-fluoro-5-vinyl-uridine (**3.6**).
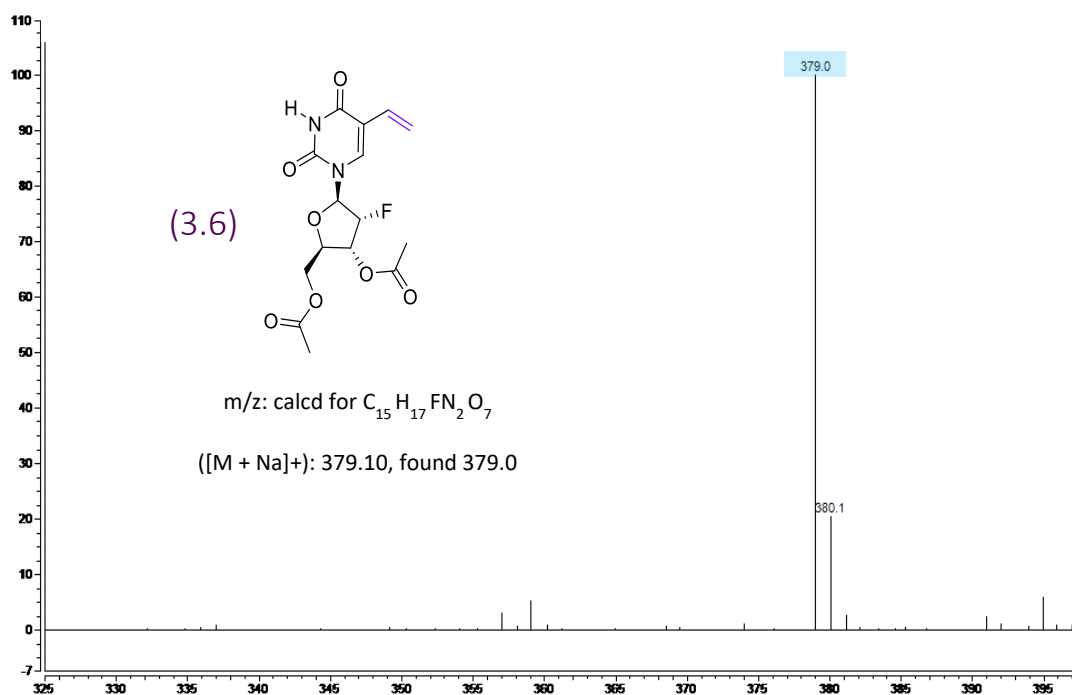
Appendix I: Figure 3.66. Mass spectra of 3',5'-di-O-acetyl-2'-desoxy-2'-fluoro-5-vinyl-uridine (**3.6**).
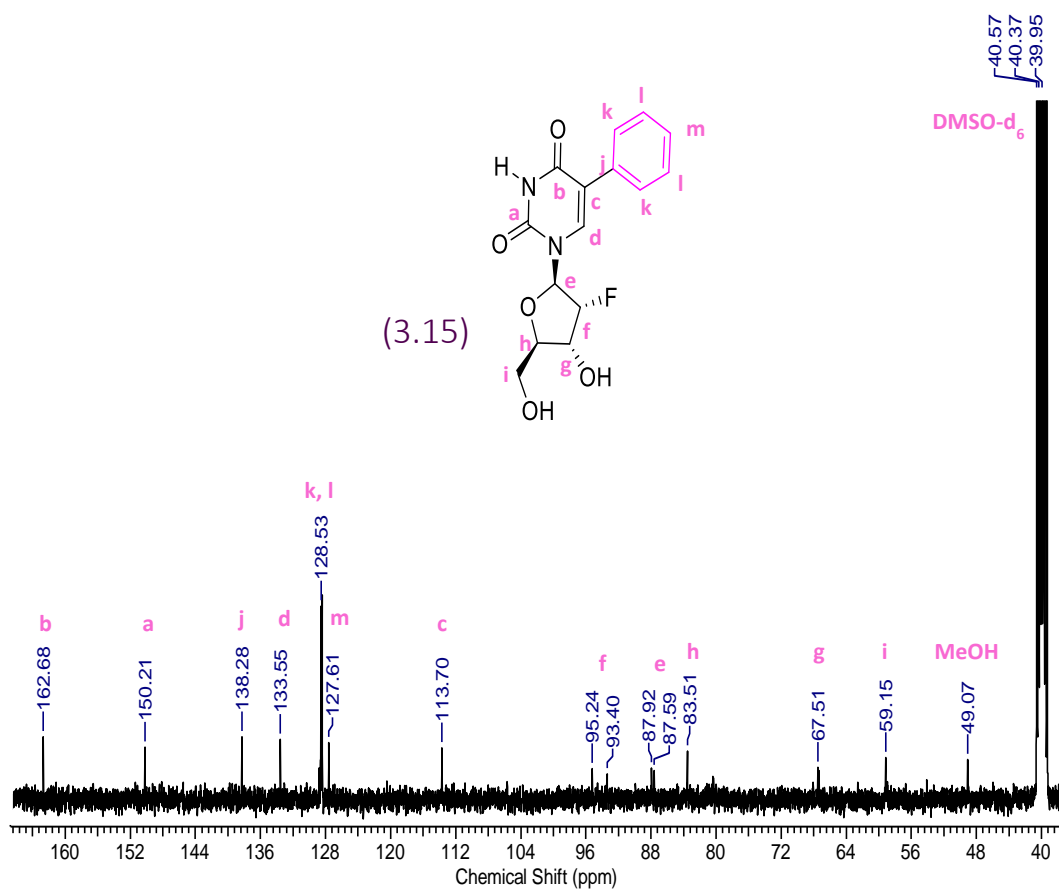


Appendix J: Figure 3.67. $^{13}$C NMR Spectra of 2'-desoxy-2'-fluoro-5-phenyl-uridine (**3.15**).
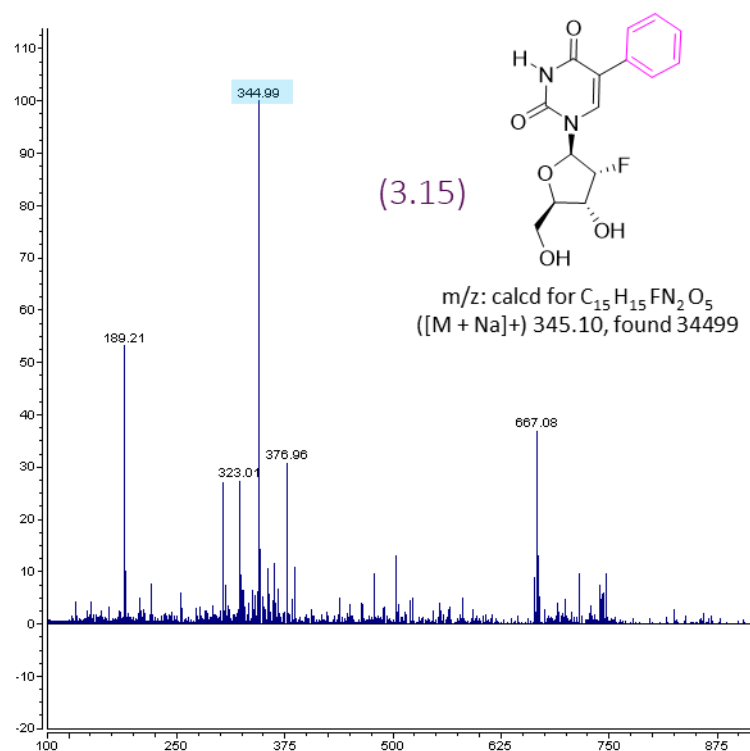
Appendix K: Figure 3.68. Mass spectra of 2'-desoxy-2'-fluoro-5-phenyl-uridine (**3.15**).
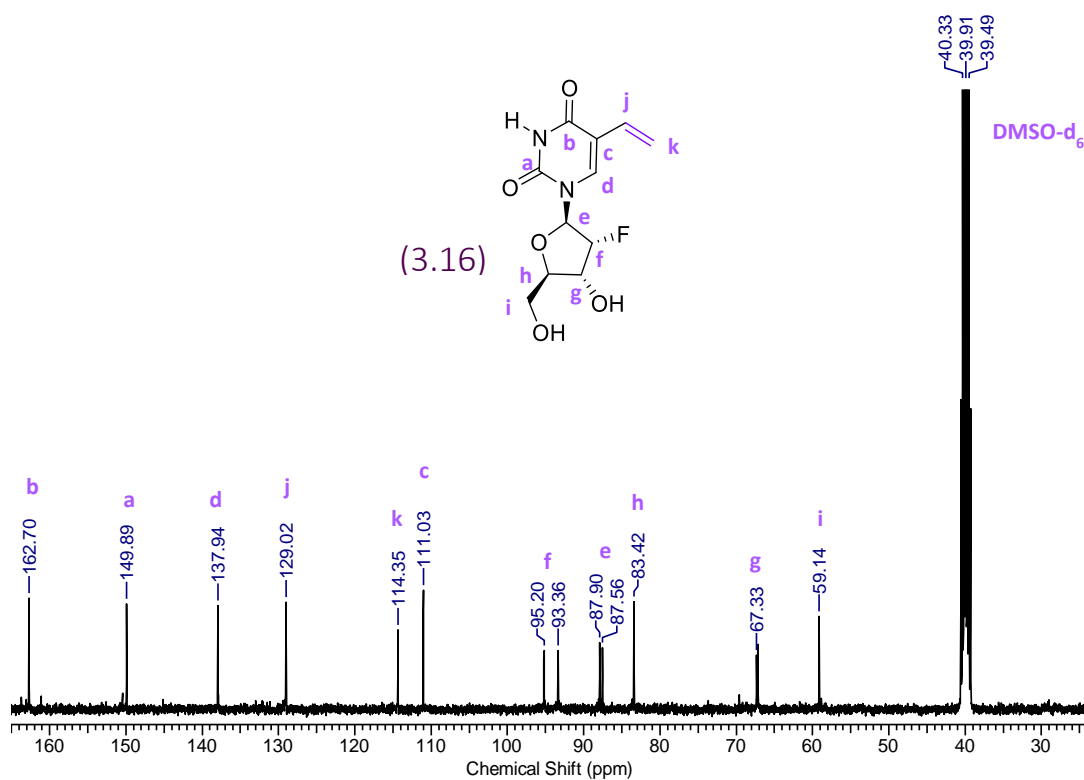


Appendix L: Figure 3.69. $^{13}$C NMR spectra of 2'-desoxy-2'-fluoro-5-vinyl-uridine (**3.16**).
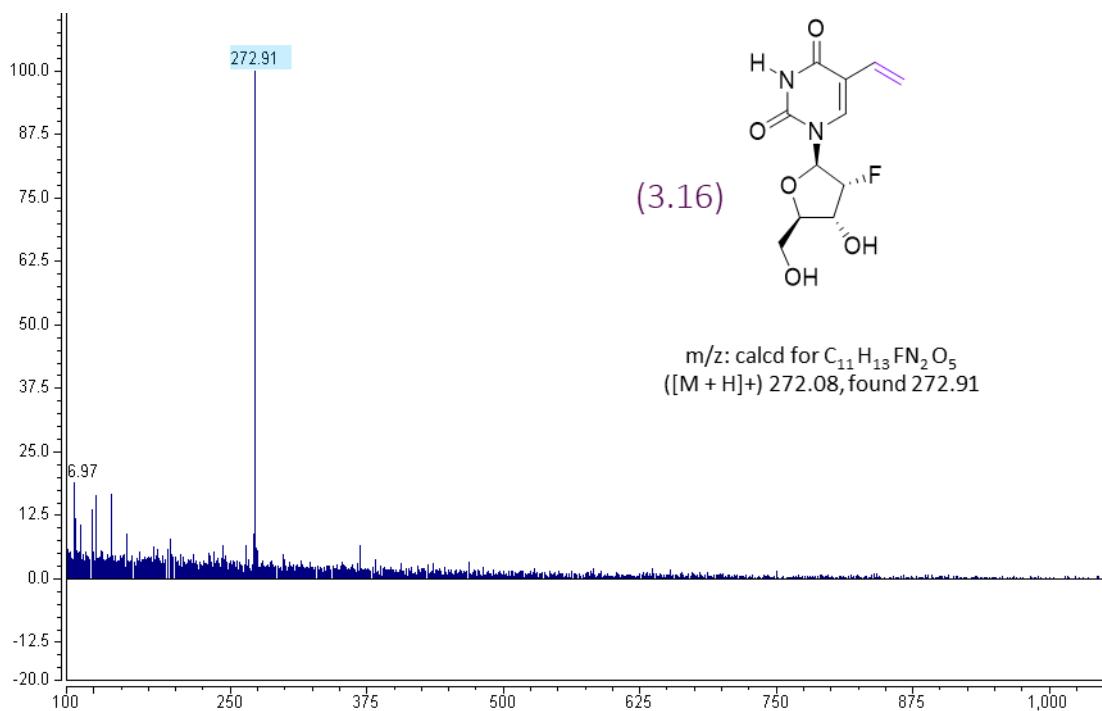
Appendix M: Figure 3.70. Mass spectra of 2'-desoxy-2'-fluoro-5-vinyl-uridine (**3.16**).
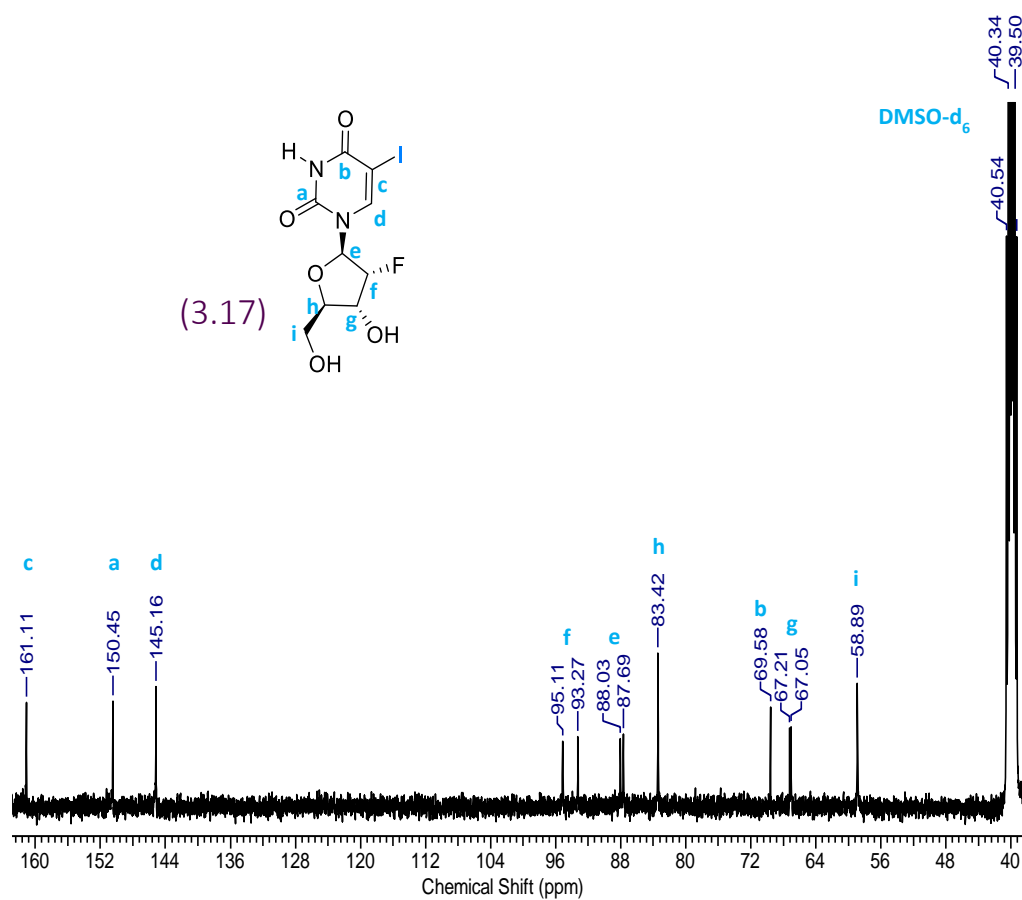


Appendix N: Figure 3.71. $^{13}$C NMR spectra of 2'-desoxy-2'-fluoro-5-iodo-uridine (**3.17**).

189

m/z: calcd for $C_9H_{10}FN_2O_5$
([M + H]+) 372.96, found 373.9
([M + Na]+) 395.10, found 394.85
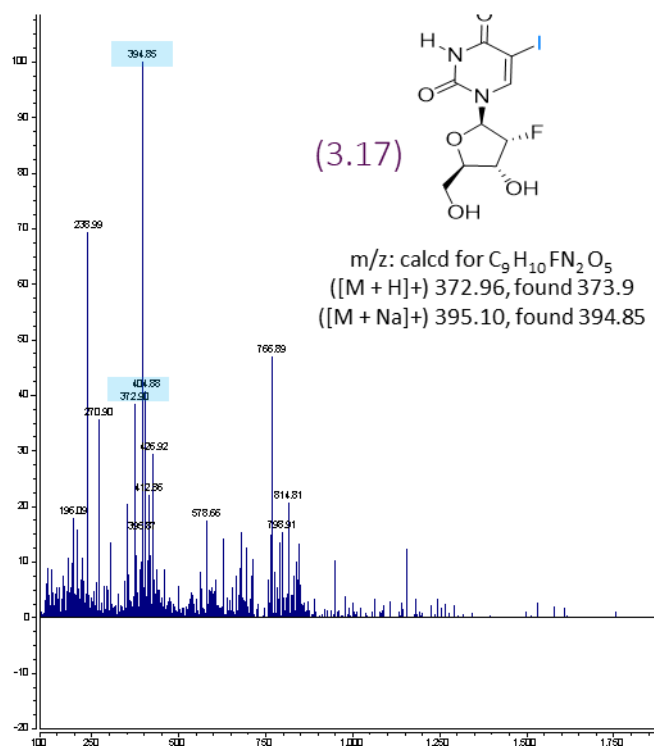
Appendix O: Figure 3.72. Mass spectra of 2'-desoxy-2'-fluoro-5-iodo-uridine (**3.17**).
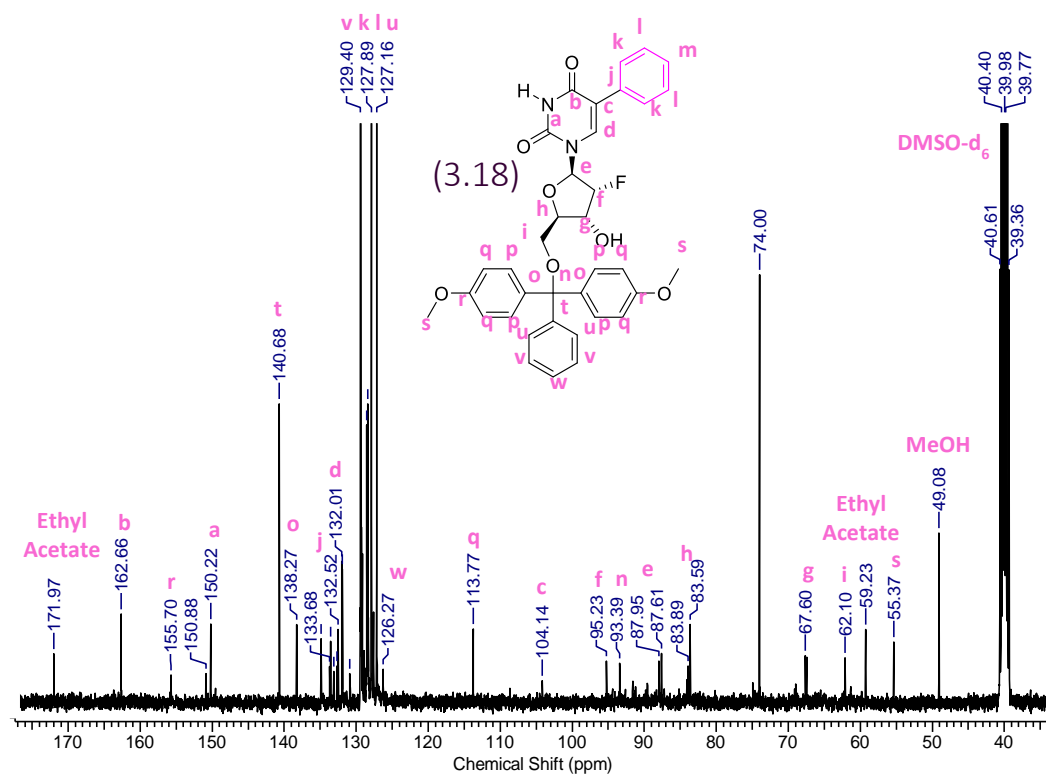


Appendix P: Figure 3.73. $^{13}$C NMR spectra of 4'4-DMT tagged 2'-desoxy-2'-fluoro-5-phenyl-uridine (**3.18**). he extra

peak at 74.00 ppm from unknown contaminate, potentially another purification step is needed.
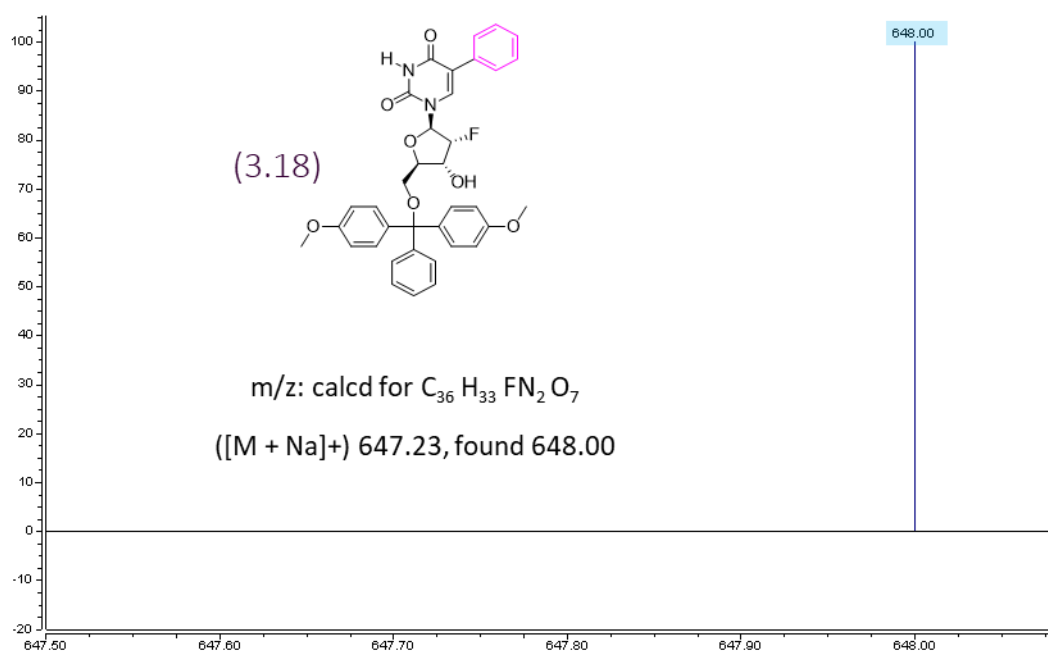
Appendix Q: Figure 3.74. Mass spectra of 4'4-DMT tagged 2'-desoxy-2'-fluoro-5-phenyl-uridine (**3.18**).
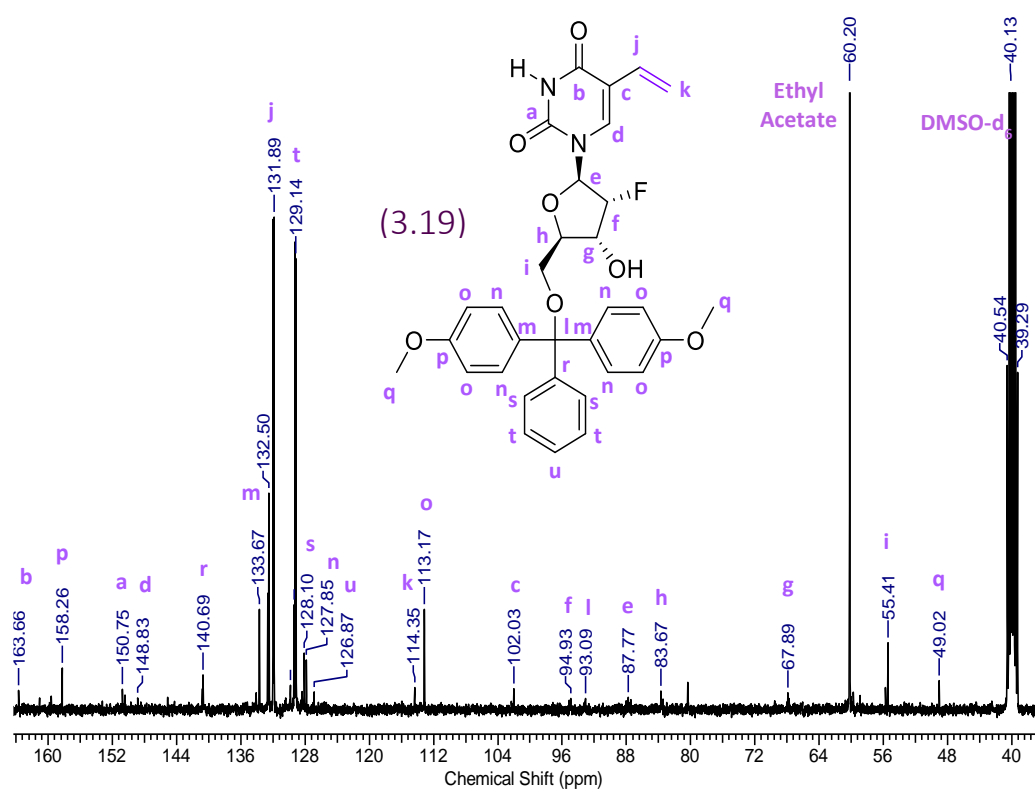


Appendix R: Figure 3.75. $^{13}$C NMR spectra of 4'4-DMT tagged 2'-desoxy-2'-fluoro-5-vinyl-uridine (**3.19**).
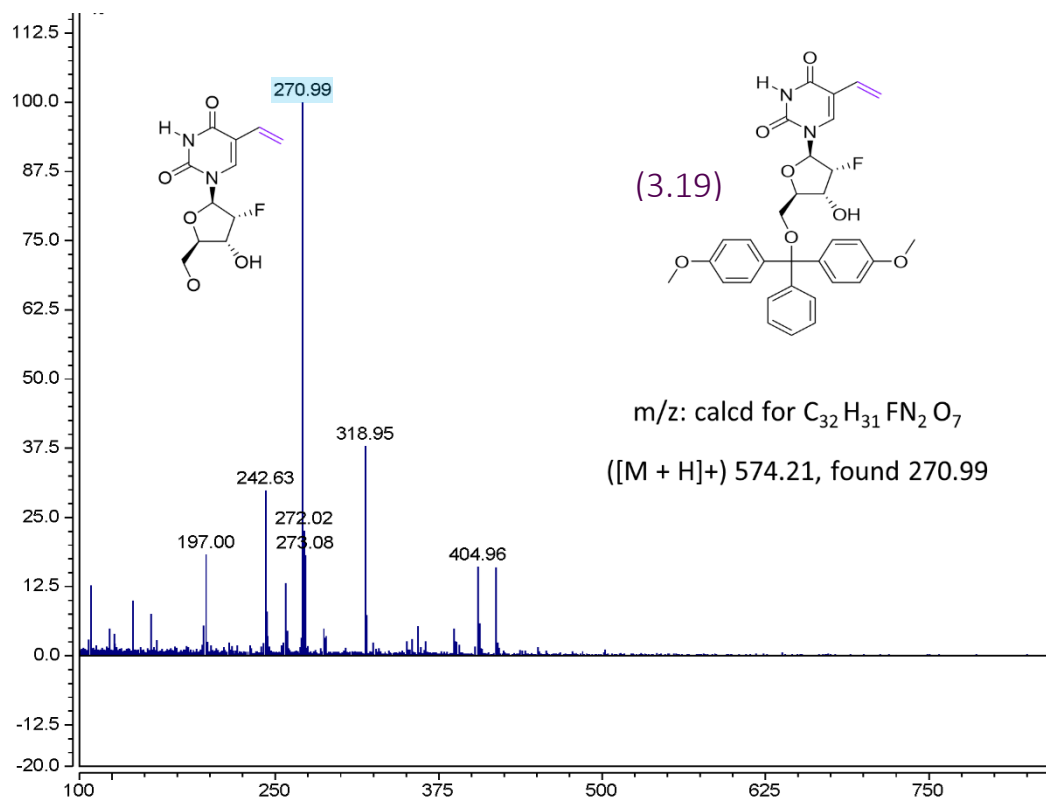
Appendix S: Figure 3.76. Mass spectra of 4'4-DMT tagged 2'-desoxy-2'-fluoro-5-vinyl-uridine (**3.19**).
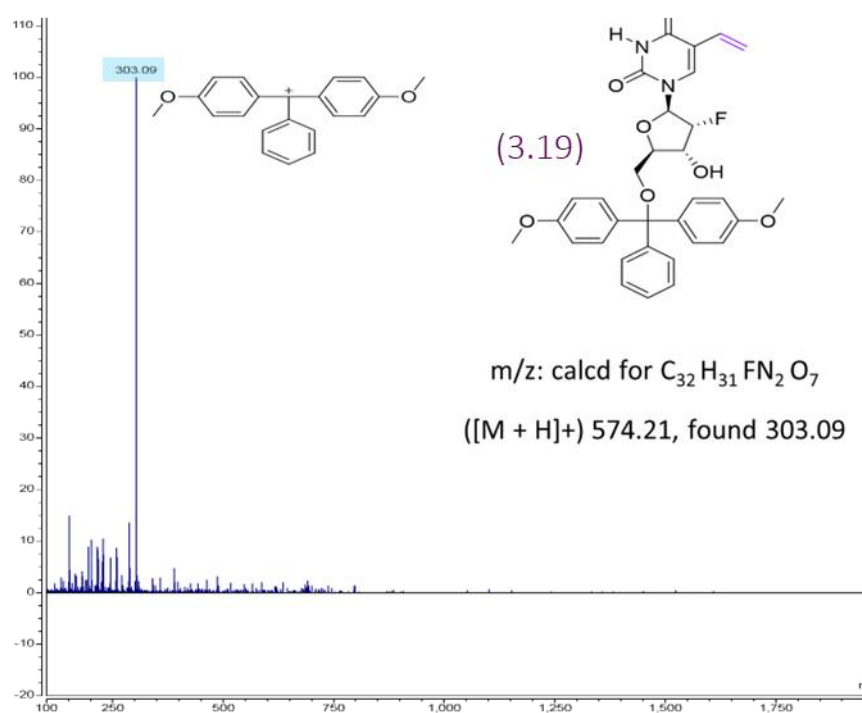


Appendix T: Figure 3.77. Mass spectra of 4'4-DMT tagged 2'-desoxy-2'-fluoro-5-vinyl-uridine (**3.19**).
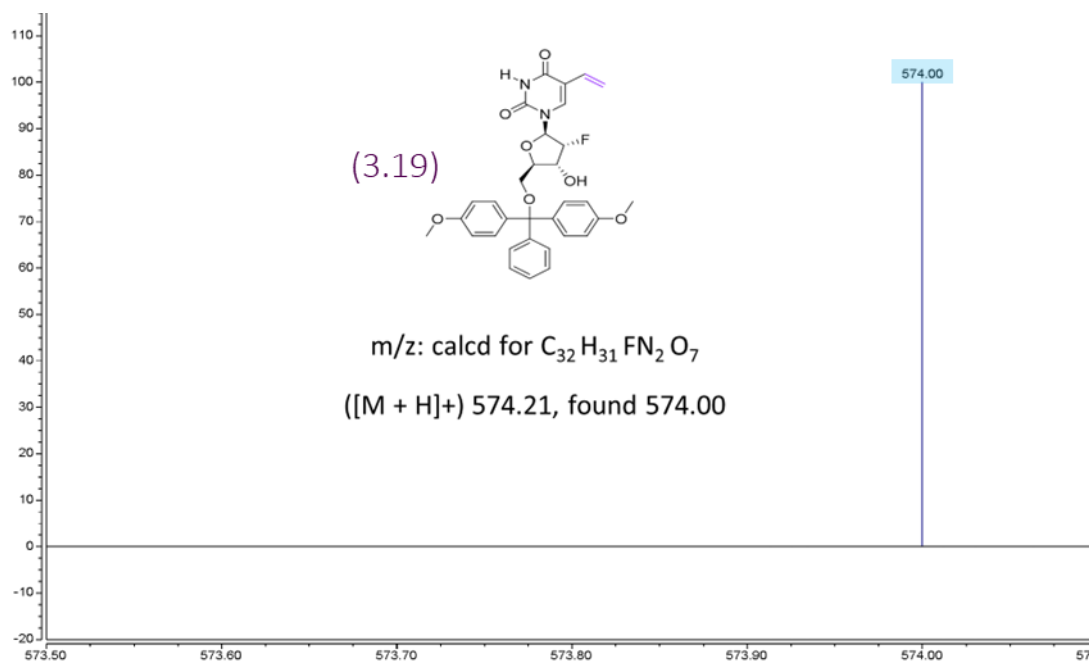
Appendix U: Figure 3.78. Mass spectra of 4'4-DMT tagged 2'-desoxy-2'-fluoro-5-vinyl-uridine (**3.19**).

m/z: calcd for $C_{32}H_{31}FN_2O_7$

([M + H]+) 574.21, found 574.00
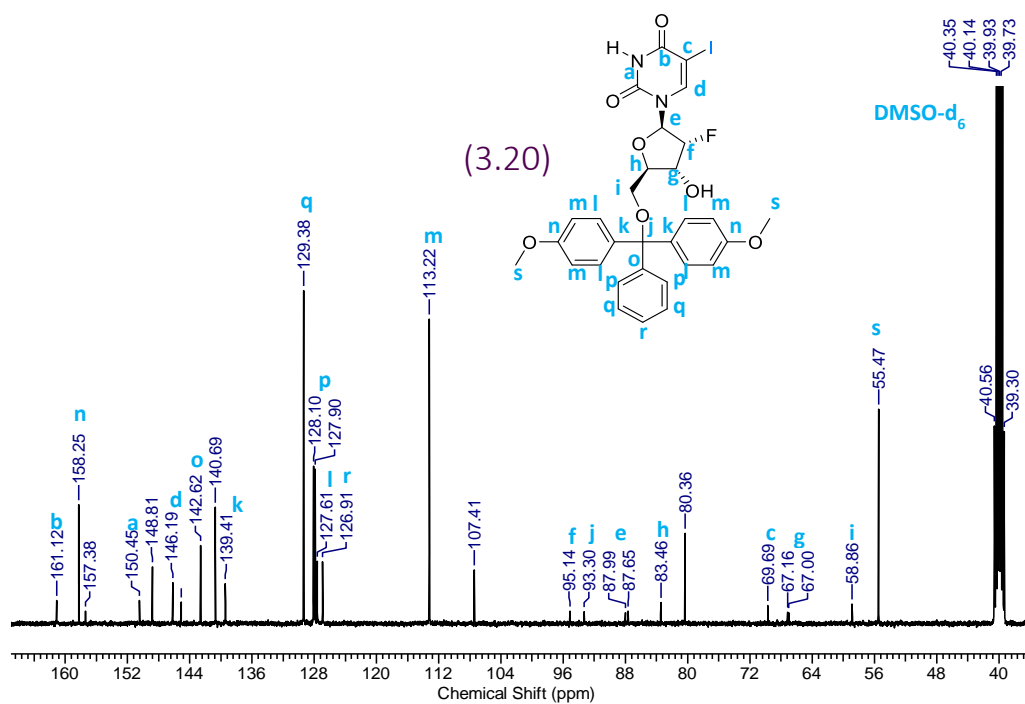


Appendix V: Figure 3.79. ^{13}C NMR spectra of 4'4-DMT tagged 2'-desoxy-2'-fluoro-5-iodo-uridine (**3.20**). There are

extra peaks at 80.36, 107.41, 140.69, 148.81 ppm which are from unknown contaminant. An extra purification step

may be needed for this compound.
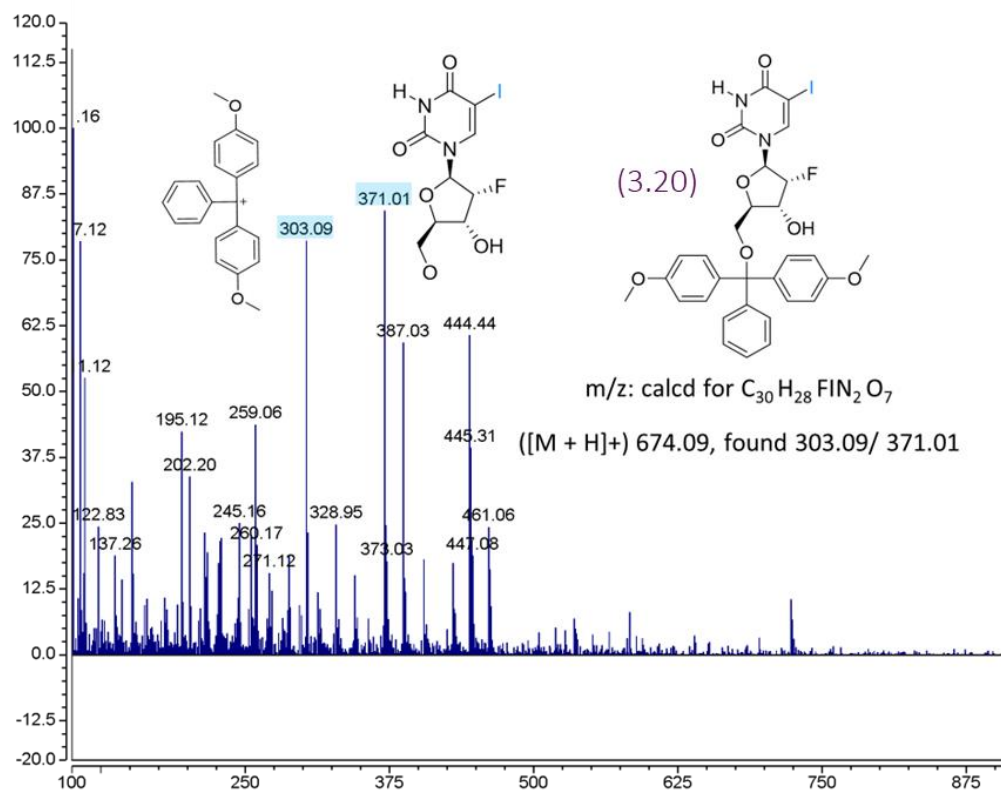
Appendix W: Figure 3.80. Mass spectra of 4'4-DMT tagged 2'-desoxy-2'-fluoro-5-iodo-uridine (**3.20**).
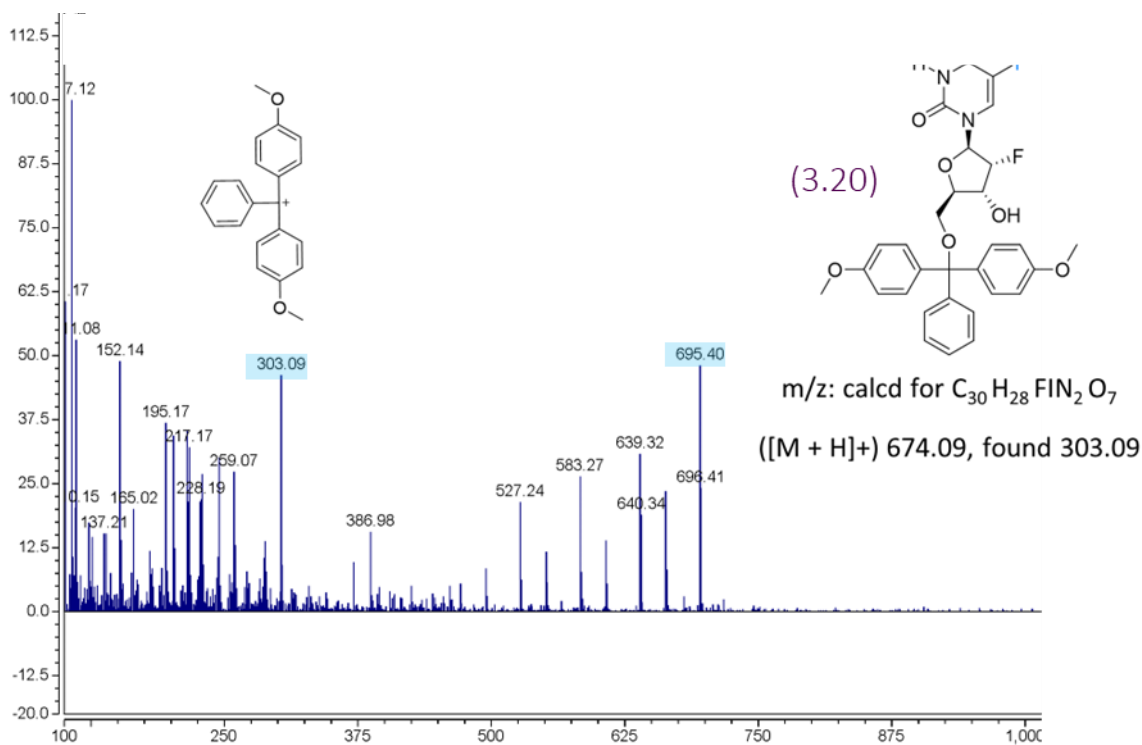


Appendix X: Figure 3.81. Mass spectra of 4'4-DMT tagged 2'-desoxy-2'-fluoro-5-iodo-uridine (**3.20**).
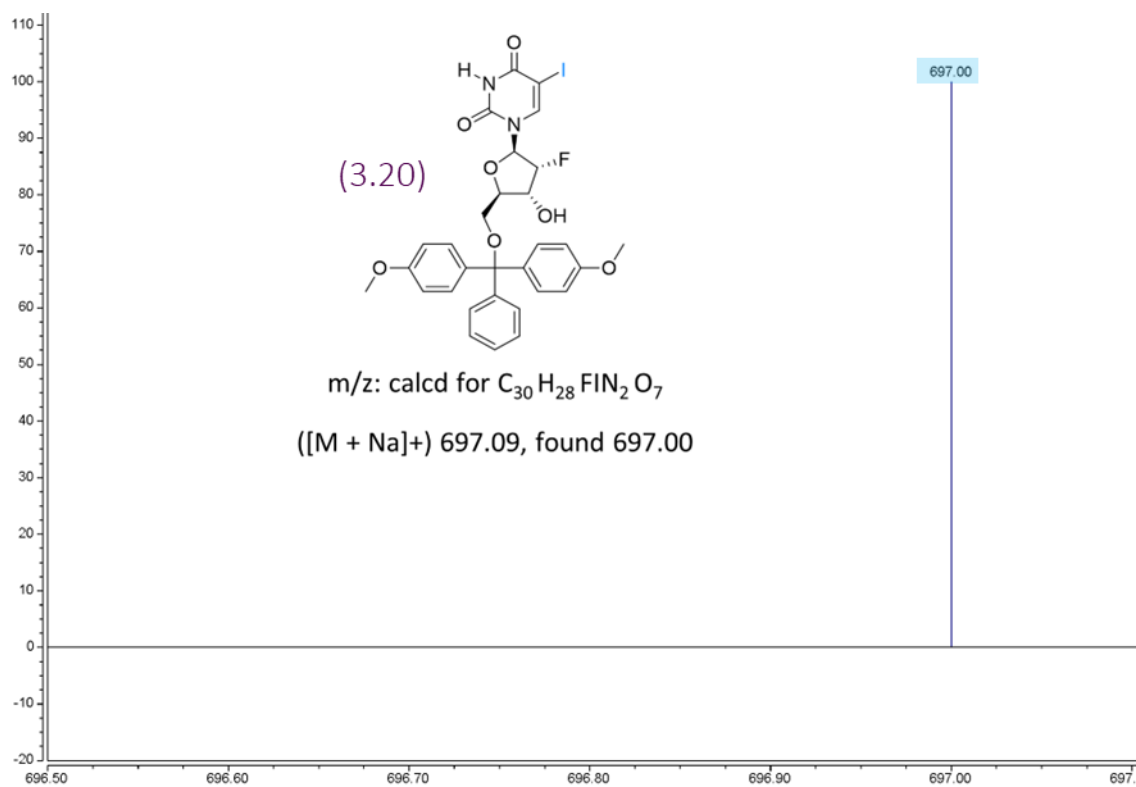
Appendix Y: Figure 3.82. Mass spectra of 4'4-DMT tagged 2'-desoxy-2'-fluoro-5-iodo-uridine (**3.20**).
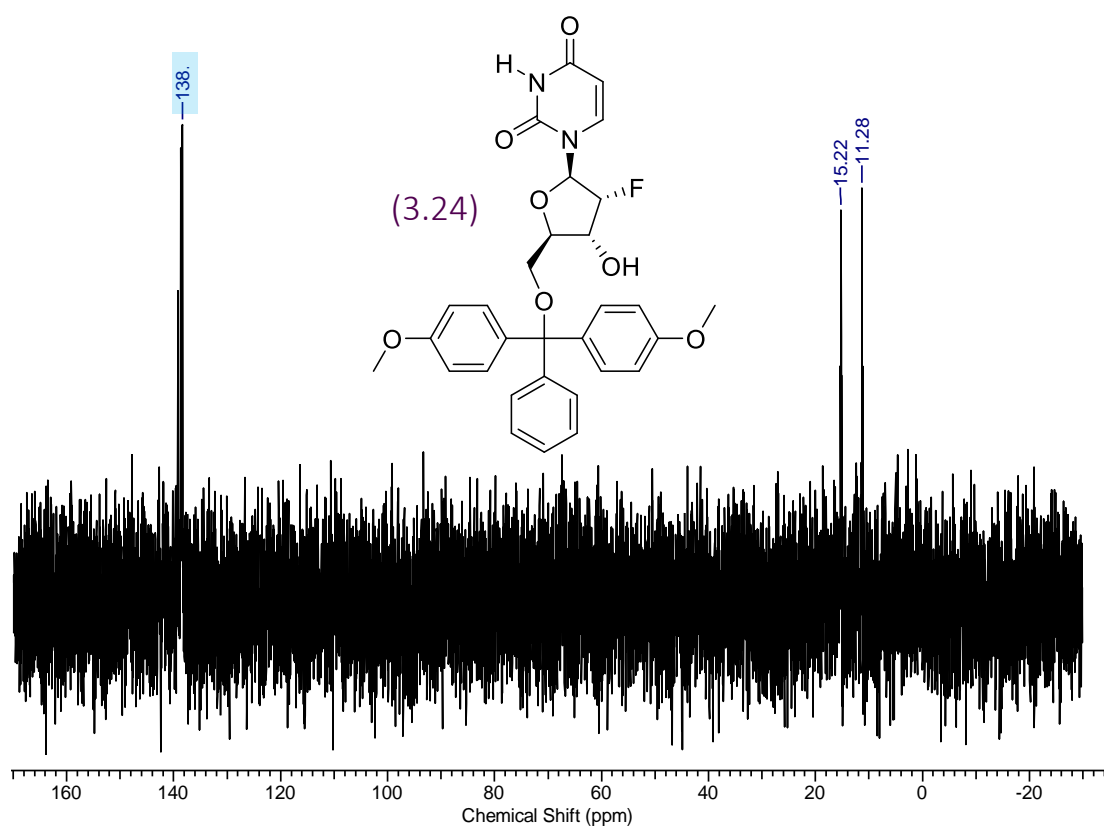


Appendix Z: Figure 3.83. $^{31}$P NMR spectra of 4'4-DMT tagged 2'-desoxy-2'-fluoro-uridine phosphoramidite (**3.24**).

# Chapter 4

## Synthesis of the Aptamer Library and Control Aptamers

# Chapter 4

## 4.1. Introduction

### 4.1.1. Aptamer Libraries

#### 4.1.1.1. Adding Diversity to Aptamer libraries

SELEX is a powerful tool for identifying aptamers for a wide range of molecular targets.[438] The effectiveness of the aptamers in research, therapeutic applications and diagnostics has been inhibited by the lack of chemical diversity of nucleic acid libraries. This limitation is partially compensated for by the large random libraries that are available to be screened. Numerous efforts have been made to increase the functional diversity of nucleic acids. This has been tried by modifying nucleobases or the sugar-phosphate backbone, aiming to generate more diversity in the aptamers to improved binding properties and metabolic stability. One approach by Benner *et al.* has been to modify the base pairing, giving a unique hydrogen bonding pattern.[439] This method went from the normal 4 letter library to a 6 letter library. These libraries fit with the PCR, amplification and sequencing of SELEX. This has resulted in the successful identification of aptamers for cell surface targets.[439] Even though the complexity of the aptamer library has been increased it is still largely similar to the natural bases.

Libraries which replace ribose with other sugars such as XNAs, locked nucleic acids (LNA),[440] L-enantiomer DNA/RNA (L-DNA)[334] and other modifications such as 2'-amino, 2'-fluoro, and 2'-O-methyl[441] have been of significant interest because of their aid in resisting nuclease degradation (figure 4.1). However, none of these have provided a wide range of functional group diversity,[438] which needs synthetic development. Nucleobase functional groups that are similar to amino acid side chains or small molecules have been introduced at either the 5-position of pyrimidines

(cytosine and thymine) or the 7- or 8-position of purines (adenine and guanine) to enhance ligand binding properties or catalytic activity.[442] Of the functional groups tested, hydrophobic groups that are similar to side chains in proteins have been proven to be the most effective modifications for a wide range of target proteins.[443] These modifications introduced at the 5' position of deoxyuridine[444] have allowed for the additional demand for kinetic stability of the complexes during selections and identification of SOMAmer reagents for over 3,000 proteins.[445]
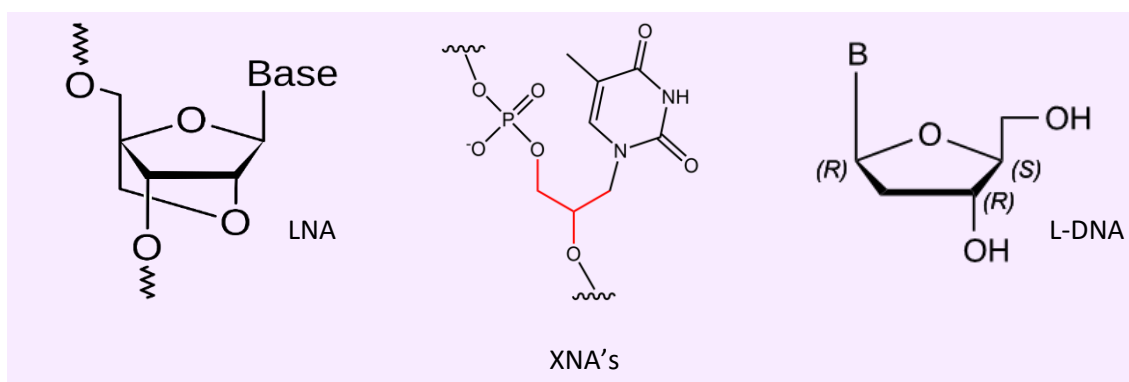


Figure 4.1. Modified Nucleic Acids XNA's: (LNA, L-DNA).

This project will make a modified aptamer library using phosphoramidite chemistry and the split and mix method. It will not follow the normal flow of SELEX; as selection against the EGFR target will be done using flow cytometry-based sorting and hit aptamer identification by MS/MS. The modifications will be on 5' position of the uridines (compounds made in chapter 3) and 2'F modification on the cytosine and uridines, adding diversity to the aptamer library.

## 4.1.2. Split and Mix Method
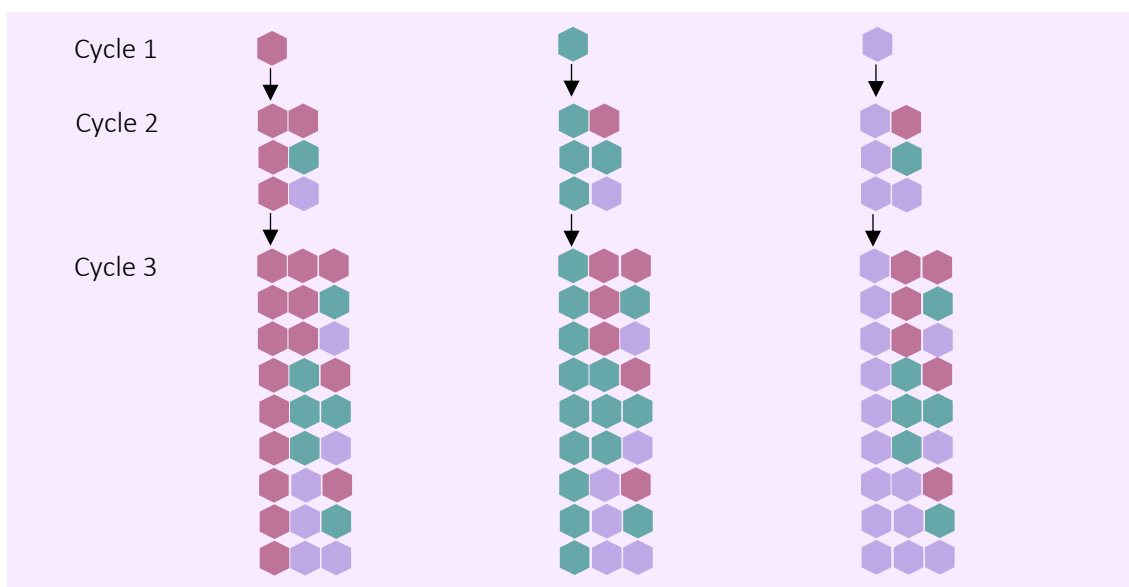


Figure 4.2. Split-Mix Synthesis.

Synthesis of OBOC libraries is done by split synthesis, also known as the 'split and mix synthesis' which is a solid phase synthesis. It works by having a combinatorial library, for example a peptide library which was synthesised using 20 amino acids. The bead support is then divided into 20 equal portions. After that a different amino acid is coupled to each portion. Finally, these portions are all mixed together. This can be repeated in a cycle and the chain will be elongated.[446] The way that a one-bead-one-compound library is synthesised is often down to the chosen polymeric carrier. A beaded polymer needs to reach certain standards depending on the screening and synthetic approach. It is very important for the substitution and the size to be the same for one-bead-one-compound libraries. It is also important to have a resin with good resistance to the creation of clusters as this could prevent statistical distribution. It is very important that the resin has the ability to swell in both organic and aqueous media when bead binding assays are used for screening.[161]

### 4.1.3. Computational Modelling

Computational modelling has gained an increasingly significant role in biomolecular and chemical fields over the past decades. This has been helped by developments in terms of software and methodology as well as technological advances in computational hardware. Today we are able to screen virtual libraries of millions of potentially therapeutic compounds searching for candidate inhibitors and targets, run simulation of large biomolecular complexes or predict protein structures with comparable accuracy to high resolution X-ray crystallography.[447] Being able to visualise the structure and shapes of compounds and therapeutic agents is an extremely useful tool for drug discovery.

Initially, the focus was on force-field based methodologies for studying the structures, dynamics and interactions of biomolecules. This was followed by the development of quantum chemical techniques suitable for studying large systems, density functional theory being the key approach for investigating enzymatic mechanisms or properties and reactions of small biologically relevant molecules.[447]

Computational quantum chemistry: molecular orbitals and their electrons, allow us to calculate chemical and physical properties that directly depend on the electron distribution. This can reach all the way to coarse-grained molecular dynamics simulations, where groups of atoms interacting by laws of Newtonian mechanics will provide valued insights into the complexity of biological processes at the cellular level.[447] The potential energy is calculated by adding up the energy terms that describe interactions between bonded atoms and terms that describe the non-bonded interactions such as electrostatic interactions and Van der Waals forces (figure 4.3).
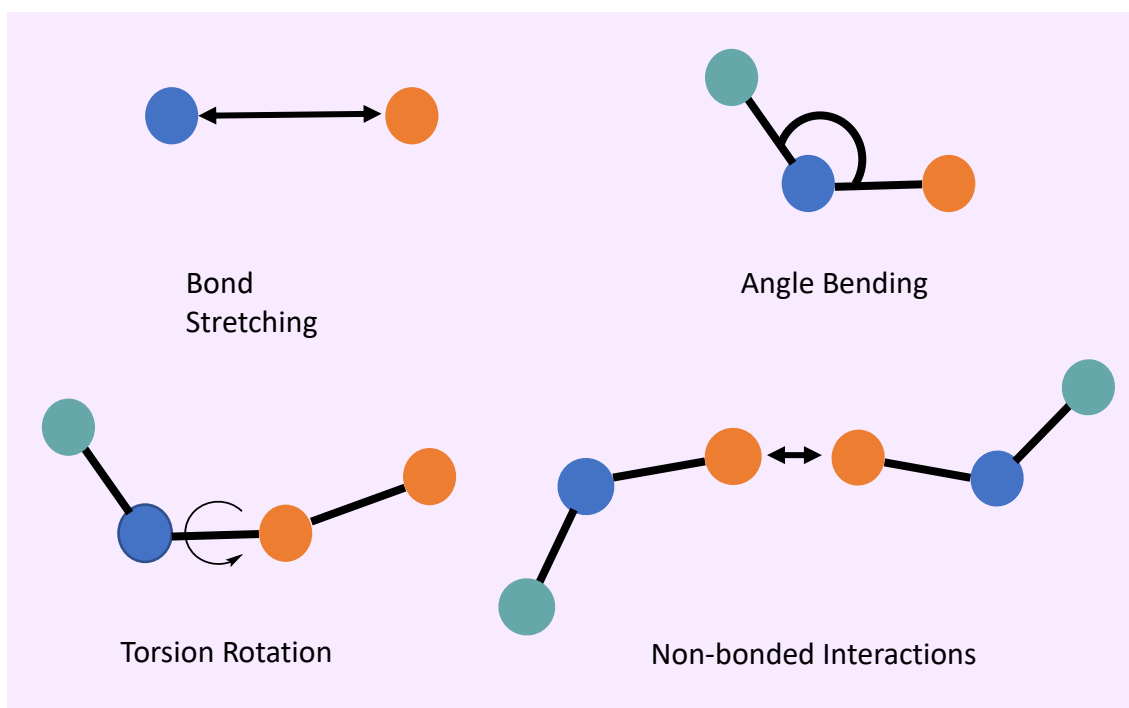
Figure 4.3. In molecular mechanics, the atoms are represented as charged spheres, which have bonded (bond stretch, angle bend and torsional angle rotation) and non-bonded interactions (Van der Waals and electrostatics).

There are a number of computational modelling software programs available, having different degrees of complexity, and oriented to treat different kinds of systems. The most commonly used ones are AMBER[448,449] CHARMM[450] and GROMOS[451] which are designed for biological macromolecules.

### 4.1.3.1. Computational Modelling of Nucleic Acids

The conformational space of DNA/RNA is diverse and dynamic and has the ability to change depending on the physicochemical properties of the surrounding environments and with interactions with other molecules. Its geometry can be described in terms of helical parameters, (pitch and diameter of the helix), groove parameters (depth and width), furanose ring conformation, six torsional angles of the backbone, rotational (tip and inclination) and translational base pair parameters, six intra-base parameters (buckle, propeller, opening, shear, stretch, and stagger), and six inter-base parameters (tilt, roll and twist, shift, slide, and rise).[447]

### 4.1.3.2. Computational Modelling for Drug Discovery

Efficient and effective drug discovery is still the biggest challenge facing the medical scientific community. The challenges faced in the identification, development and clinical trials associated with novel drug discovery are not only limited to the complexity of the task but also by the significant time and financial requirements. It has been reported that the total cost of drug delivery from bench to drug approval averages 2.6 billion USD and spans 14 years.[452] These challenges highlight the need for new techniques for rapid screening and discovery of therapeutic candidates at low cost. A technique that is helping this is molecular modelling, with molecular docking, molecular dynamics (MD) simulation and ADMET modelling representing the most widely used specific techniques.[453,454,455] Molecular docking-based virtual screening will identify hit compounds with high affinity[456] for their target, however, it often lacks simulation of receptor flexibilities. Conformation of proteins plays a vital role in the biomolecular recognition of ligands and so their flexibility in a simulation environment represents a more accurate description of the biological system. Some docking software packages have been developed to address this, however a large number of these programs do not factor in the full flexibility of the receptors. MD simulations are able to take on this task to determine the time-dependent dynamics of protein-ligand interactions. MD simulations treat ligands, proteins, water molecules and ions as particles interacting with one another by the integration of force fields that have been derived on the basis of Newton's classical law of motion. The simulations allow for the inspection of the stability of ligands in the active pocket of the receptor targets. This is crucial to validating the results predicted through the molecular docking-based virtual screening process.[457]

ADMET modelling is capable of predicting the clinical success of drugs. The pharmaceutical industry has a challenge in coordinating the optimisation of the combination of pharmacokinetic and pharmacodynamic properties that are driven by characteristics such as

absorption, distribution, metabolism, excretion and toxicity. Because of this there has been a big effort to create tools that can predict pharmacokinetics and pharmacodynamic endpoints during the optimisation process such as minimising animal models and having better targeting clinical trials.[458] ADMET prediction platforms can reject inappropriate compounds by simultaneously targeting numerous pharmacokinetic characteristics. This lowers the amount of synthesis and late-stage trial failures.

All of these techniques in drug discovery are hugely important. Although there are limitations, pairing them with experimental data is very powerful and can reduce the amount of both time and money spent on drug discovery.

### 4.1.4. Chapter Aims

The aim of this chapter is to optimise the synthesis of the aptamers before synthesising the control aptamers and fully modified aptamers (**MinE07-Biotin, MinE07M3-Biotin, MinE07M6-Biotin, MinE07-U-Ph-Biotin, MinE07-U-Vi-Biotin,** and **MinE07-U-I-Biotin**) to be used as a comparison for aptamers in the MinE07 aptamer library. After this optimisation, the aptamer library will be synthesised on the solid support –OH modified TentaGel microspheres, using the split and mix method. Computational modelling will be used to give visualization of how the modifications can affect the shape and folding of the MinE07 aptamer.

## 4.2. Materials and Instrumentation

### 4.2.1. Materials

Synthesiser reagents: oxidizer (0.02M iodine, 20% pyridine), Cap A Mix (THF/Pyridine/acetic anhydride 8:1:1), Cap B Mix (10% methylimidazole in THF), d-block (3% trimethylamine in DCM) and ETT Activator (0.25 M, 5-ethylthio-1H-tetrazole in acetonitrile) were purchased from Link Technologies. N2-Acetyl-2'-O-tert-butyldimethylsilyl-5'-O-DMT-guanosine 3'-CE

phosphoramidite, N6-Benzoyl-2'-O-tert-butyldimethylsilyl-5'-O-DMT-adenosine 3'-CE phosphoramidite, N4-Acetyl-2'-deoxy-5'-O-DMT-2'-fluorocytidine 3'-CE phosphoramidite and 2'-Deoxy-5'-O-DMT-2'-fluorouridine 3'-CE phosphoramidite were purchased from CarboSynth. TentaGel® M NH₂ Monosized Amino TentaGel Microspheres were purchased from Rapp Polymere.

## 4.2.2. Instrumentation

### 4.2.2.1. Expedite™ 8909 DNA Synthesiser

Specification can be found in Chapter 2 section 2.2.2.1 Expedite™ 8909 DNA Synthesiser

### 4.2.2.2. Mass Spectrometer

Specification can be found in Chapter 3 section 3.2.2.3 Mass Spectrometer

### 4.2.2.3. UV-Visible Absorption

Specification can be found in Chapter 2 section 2.2.2.5 UV-Visible Absorption Spectra

# 4.3. Experimental Methods

### 4.3.1. Set up for the synthesis of All MinE07 Aptamers

The phosphoramidite samples were dissolved in 20 mL (20 mL for 1 g of sample) acetonitrile (DCM for Tentagel beads), put in bottles and screwed into the synthesiser lines. The other reagents put on to the machine: oxidizer (0.02M iodine, 20% pyridine), Cap A Mix (THF/Pyridine/acetic anhydride 8:1:1), Cap B Mix (10% methylimidazole in THF), deblock (3% trimethylamine in DCM) and ETT activator solution (0.25 M, 5-ethylthio-1H-tetrazole in acetonitrile). A leak test is run to check nitrogen is not leaking from the lines. If passed, the lines are then flushed with the new reagents added to them. The beads are added to the column, which is then fitted onto the synthesiser, which is then flushed with acetonitrile several times.

The sequence and protocol are then selected using Validate XP connected to the Expedite™ 8909 DNA Synthesiser The aptamer sequences are then run and monitored using the trityl monitor.

### 4.3.2. Synthesis of MinE07 Aptamer

Universal UnyLinker support (0.021g) was added to a synthesiser column. See experimental method 4.3.1 for the setup of the synthesiser. The sequence uploaded on to the synthesiser was: 5'-rGrGrAfCrGrGrAfUfUfUrArAfUfCrGfCfCrGfUrArGrArArArArGfCrAfUrGfUfCrArArArGfCfC rGrGrArAfCfCrGfUfCfC-3'. This was run to completion.

### 4.3.3. 2D Predicted secondary structure of MinE07 aptamer

ViennaRNA (rna.tbi.univie.ac.at/) was used to model a 2D predicted secondary structure of MinE07 aptamer, using the thermodynamic structure prediction and the RNAfold web server giving a minimum free energy of -9.30kcal/mol. A dot and bracket sequence was generated to be used to create a PDB file.  ((((((.........(((.......((........)))))..))))))

### 4.3.4. 3D Predicted secondary structure of MinE07 aptamer

The dot and bracket sequence was used to generate a PDB file to create a 3D structure  of the MinE07 aptamer using RNA composer (rnacomposer.cs.put.poznan.pl/). The PDB file was opened in Wordpad and all the 2'OH atoms were deleted from uridine and cytidine (Figure. 4.4).

| ATOM | 285 | C2' | U A | 9 | -2.379 -17.851 -16.606 | 1.00 | 0.00 | C |
| ATOM | 286 | H2' | U A | 9 | -2.042 -16.833 -16.402 | 1.00 | 0.00 | H |
| ATOM | 287 | O2' | U A | 9 | -1.492 -18.832 -16.101 | 1.00 | 0.00 | O |
| ATOM | 288 | HO2' | U A | 9 | -1.937 -19.252 -15.364 | 1.00 | 0.00 | H |

Figure 4.4. PBD File of MinE07 aptamer before editing.

The O2' lines were then replaced with F2' (Figure 4.5)

```
ATOM  283  C2' U  A  9    -2.332 -17.991 -16.814  1.00  0.00        C
ATOM  284  H2' U  A  9    -1.836 -17.068 -16.520  1.00  0.00        H
ATOM  285  F2' U  A  9    -1.513 -19.038 -16.468  1.00  0.00        F
```

Figure 4.5. PBD File of MinE07 aptamer after editing.

This PDB can be visualised using the cross platform molecular builder and editor program, Avogadro (avogadro.cc). The geometry was then optimised using the optimise geometry option in Avogadro with force field MMFF94, which is in a vacuum.

**4.3.5. Synthesis of Biotin-MinE07 Aptamer**

Universal UnyLinker support (0.021 g) was added to a synthesiser column. See experimental method 4.3.1 for the setup of the synthesiser. The sequence uploaded on to the synthesiser was: 5'-Biotin-rGrGrAfCrGrGrAfUfUfUrArAfUfCrGfCfCrGfUrArGrArArArArGfCrAfUrGfUfCrArArA rGfCfCrGrGrArAfCfCrGfUfCfC-3'. This was run to completion.

**4.3.6. Synthesis of Biotin-MinE07M3 Aptamer**

Universal UnyLinker support (0.021 g) was added to a synthesiser column. See experimental method 4.3.1 for the setup of the synthesiser. The sequence uploaded on to the synthesiser was: 5'-Biotin-rGrGrAfCrGrGrAfUfUfUrArAfUfCrGfCfCrGfUrArGrArArArArGfCrAfUrGfUfCrArArA rGfCfCrGrGrArAfCfCrG-3'. This was run to completion.

**4.3.7. 2D Predicted secondary structure of MinE07M3 aptamer**

ViennaRNA (rna.tbi.univie.ac.at/) was used to model a 2D predicted secondary structure of the MinE07M3 aptamer, using the thermodynamic structure prediction and the RNAfold web server giving a minimum free energy of -3.80kcal/mol. A dot and bracket sequence was generated to be used to create a PDB file.  ..((((.........))))......((........))........

### 4.3.8. 3D Predicted secondary structure of MinE07M3 aptamer

The dot and bracket sequence was used to generate a PDB file to create a 3D structure using the RNA composer (rnacomposer.cs.put.poznan.pl/). The PDB file was opened in Wordpad and all the 2'OH atoms are deleted from uridine and cytidine (Figure 4.6)

```
ATOM    285  C2'  U A  9    -2.379 -17.851 -16.606  1.00  0.00    C
ATOM    286  H2'  U A  9    -2.042 -16.833 -16.402  1.00  0.00    H
ATOM    287  O2'  U A  9    -1.492 -18.832 -16.101  1.00  0.00    O
ATOM    288  HO2' U A  9    -1.937 -19.252 -15.364  1.00  0.00    H
```

Figure 4.6. PBD File of MinE07M3 aptamer before editing.

The O2' lines were then replaced with F2' (Figure 4.7).

```
ATOM    283  C2'  U  A  9    -2.332 -17.991 -16.814  1.00  0.00    C
ATOM    284  H2'  U  A  9    -1.836 -17.068 -16.520  1.00  0.00    H
ATOM    285  F2'  U  A  9    -1.513 -19.038 -16.468  1.00  0.00    F
```

Figure 4.7. PBD File of MinE07M3 aptamer after editing.

This PDB can be visualised using the cross platform molecular builder and editor program, Avogadro (avogadro.cc). The geometry was then optimised using the optimise geometry option in Avogadro with force field MMFF94, which is in a vacuum.

### 4.3.9. Synthesis of Biotin-MinE07M6 Aptamer

Universal UnyLinker support (0.021 g) was added to a synthesiser column. See experimental method 4.3.1 for the setup of the synthesiser. The sequence uploaded on to the synthesiser was: 5'-Biotin-rGrGrAfCrGrGrAfUfUfUrArAfUfCrGfCfCrGfUrArGrArArArArGfCrAfUrGfUfCrArArA rGfCfCrGrGrArA-3'. This was run to completion.

### 4.3.10. 2D Predicted secondary structure of MinE07M6 aptamer

ViennaRNA (rna.tbi.univie.ac.at/) was used to model a 2D predicted secondary structure of the MinE07M6 aptamer, using the thermodynamic structure prediction and the RNAfold web server

giving a minimum free energy of -3.80kcal/mol. A dot and bracket sequence was generated to be used to create a PDB file. ..((((.........))))......((........)).....

**4.3.11. 3D Predicted secondary structure of MinE07M6 aptamer**

The dot and bracket sequence was used to generate a PDB file to create 3D structure using RNA composer (rnacomposer.cs.put.poznan.pl/). The PDB file was opened in wordpad and all the 2'OH atoms are deleted from uridine and cystine (Figure 4.8).



Figure 4.8. PBD File of MinE07M6 aptamer before editing.

The O2' lines were then replaced with F2' (Figure 4.9)



Figure 4.9. PBD File of MinE07M6 aptamer after editing.

This PDB can be visualised using the cross platform molecular builder and editor program, Avogadro (avogadro.cc). The geometry was then optimised using the optimise geometry option in Avogadro with force field MMFF94, which is in a vacuum.

**4.3.12. Synthesis of MinE07Lib Aptamer Library**

-OH modified TentaGel® M NH₂ Monosized Amino TentaGel Microspheres (0.0087 g) were added to a synthesiser column. This was enough to synthesis 200 copies of every possible sequence. See experimental method 4.3.1 for the setup of the synthesiser.

Round 1: 3'-PCLinker-fCfC-5'. This sequence was loaded onto the synthesiser and run on all TentaGel microspheres.

Round 2: Uridine 1: 3'-fU-5' – The Tentagel microspheres were taken out of the main column and split in 4 columns (U-Ph, U-Vi, U, U-I). A synthesis circle using compound **3.21** was performed on column U-Ph. A synthesis circle using compound **3.22** was run through column U-Vi. A synthesis circle using compound **3.23** was run through column U. A synthesis circle using unmodified uridine phosphoramidite was run through column U-I.

Round 3: 3'-rGfCfCrArArGrGfCfCrGrArArAfC-5'. This sequence was loaded onto the synthesiser and run on all TentaGel microspheres.

Round 4: Uridine 2: 3'-fU-5' – The same synthesis procedure was performed as for round 2.

Round 5: 3'-rG-5'. This sequence was loaded onto the synthesiser and run on all TentaGel microspheres.

Round 6: Uridine 3: 3'-fU-5' – The same synthesis procedure was performed as round 2.

Round 7: 3'-rAfCrGrArArArArGrA-5'. This sequence was loaded onto the synthesiser and run on all TentaGel microspheres.

Round 8: Uridine 4: 3'-fU-5' – The same synthesis procedure was performed as round 2.

Round 9: 3'-rGfCfCrGfC-5'. This sequence was loaded onto the synthesiser and run on all TentaGel microspheres.

Round 10: Uridine 5: 3'-fU-5' – The same synthesis procedure was performed as round 2.

Round 11: 3'-rArA-5'. This sequence was loaded onto the synthesiser and run on all TentaGel microspheres.

Round 12: Uridine 6: 3'-fU-5' – The same synthesis procedure was performed as round 2.

Round 13: Uridine 7: 3'-fU-5' – The same synthesis procedure was performed as round 2.

Round 14: Uridine 8: 3'-fU-5' – The same synthesis procedure was performed as round 2.

Round 15: 3'-rArGrGfCrArGrG-5'. This sequence was loaded onto the synthesiser and run on all TentaGel microspheres.

The weight of the final MinE07Lib TentaGel Microspheres was 0.0029 g.

### 4.3.14. Synthesis of MinE07-U-Ph-Biotin Aptamer

Universal UnyLinker support (0.021 g) was added to a synthesiser column. See experimental method 4.3.1 for the setup of the synthesiser. The sequence uploaded on to the synthesiser was: 5'-Biotin-rGrGrAfCrGrGrAfU(U-Ph)fU(U-Ph)fU(U-Ph)rArAfU(U-Ph)fCrGfCfCrGfU(U-Ph)rArG rArArArArGfCrAfU(U-Ph)rGfU(U-Ph)fCrArArArArGfCfCrGrGrArAfCfCrGfU(U-Ph)fCfC-3'. This was run to completion.

### 4.3.15. 3D Predicted secondary structure of MinE07-U-Ph aptamer

The geometry optimised MinE07 F2' edited PDB file was opened in Avogadro. The hydrogen atom at C(5) on all the uridine residues were deleted and a phenyl group was added. The geometry was then optimised in Avogadro with force field MMFF94. The aptamer was then visualised using Hermes by Cambridge Crystallographic Data Centre (CCCC) (https://www.ccdc.cam.ac.uk/solutions/csd-core/components/csd-hermes/).

### 4.3.16. Synthesis of MinE07-U-Vi-Biotin Aptamer

Universal UnyLinker support (0.021 g) was added to a synthesiser column. See experimental method 4.3.1 for the setup of the synthesiser. The sequence uploaded on to the synthesiser was: 5'-Biotin-rGrGrAfCrGrGrAfU(U-Vi)fU(U-Vi)fU(U-Vi)rArAfU(U-Vi)fCrGfCfCrGfU(U-Vi)rArGrAr ArArArGfCrAfU(U-Vi)rGfU(U-Vi)fCrArArArArGfCfCrGrGrArAfCfCrGfU(U-Vi)fCfC-3'. This was run to completion.

### 4.3.17. 3D Predicted secondary structure of MinE07-U-Vi aptamer

The geometry optimised MinE07 F2' edited PDB file was opened in Avogadro. The hydrogen atom at C(5) on all the uridine residues were deleted and a vinyl group was added. The geometry was then optimised in Avogadro with force field MMFF94. The aptamer was then visualised using Hermes (https://www.ccdc.cam.ac.uk/solutions/csd-core/components/csd-hermes/).

### 4.3.18. Synthesis of MinE07-U-I-Biotin Aptamer

Universal UnyLinker support (0.021 g) was added to a synthesiser column. See experimental method 4.3.1 for the setup of the synthesiser. The sequence uploaded on to the synthesiser was: 5'-Biotin-rGrGrAfCrGrGrAfU(U-I)fU(U-I)fU(U-I)rArAfU(U-I)fCrGfCfCrGfU(U-I)rArGrArArArA rGfCrAfU(U-I)rGfU(U-I)fCrArArArGfCfCrGrGrArAfCfCrGfU(U-I)fCfC-3'. This was run to completion.

### 4.3.19. 3D Predicted secondary structure of MinE07-U-I aptamer

The geometry optimised MinE07 F2' edited PDB file was opened in Avogadro. The hydrogen atom at C(5) on all the uridine residues were deleted and an iodo group was added. The geometry was then optimised in Avogadro with force field MMFF94. The aptamer was then visualised using Hermes (https://www.ccdc.cam.ac.uk/solutions/csd-core/components/csd-hermes/).

### 4.3.20. Cleavage of MinE07-Biotin, MinE07M3-Biotin, MinE07M6-Biotin, MinE07-U-Ph-Biotin, MinE07-U-Vi-Biotin and MinE07-U-I-Biotin Aptamers from Universal Unylinker Support

MinE07-Biotin, MinE07M3-Biotin, MinE07M6-Biotin, MinE07-U-Ph-Biotin, MinE07-U-Vi-Biotin and MinE07-U-I-Biotin aptamers were cleaved and deprotected. The solid support resin was removed from the column and placed into a screw-cap centrifuge tube and to this 1.5 mL of ammonia was added. The aptamers were incubated at 55 °C overnight in a stirring water bath. The aptamers are then placed in a centrifugal vacuum concentrator to remove the ammonia solution. The aptamers were then each re-suspended in autoclaved water.

### 4.3.21. Desalt of MinE07-Biotin, MinE07M3-Biotin, MinE07M6-Biotin, MinE07-U-Ph-Biotin, MinE07-U-Vi-Biotin and MinE07-U-I-Biotin Aptamers using Zetadex

Purification of the RNA aptamers was carried out by size exclusion gel chromatography using Zetadex resin (emp Biotech). A 20 mL column was prepared with 1 mL of cotton wool at the

bottom of the column. A slurry of zetadex resin and autoclaved water was prepared for use in the column.  1 mL of aptamer was loaded onto the column and the column was flushed with deionised water. 1 mL fractions were collected. This was repeated for each aptamer, with a new column each time.  The fractions were run on the Nanodrop UV-Vis spectrophotometer. The fractions containing an RNA signal were pooled. dried down and re-suspended in reverse-osmosis purified water and stored in the freezer (-20°C).

### 4.3.22. PAGE-based analysis of Aptamers

TBE denaturing polyacrylamide gels were prepared from a denaturing 20% polyacrylamide stock solution and TBE buffer, with the final concentrations of polyacrylamide at 15%.  The gels were polymerised by addition of 5 μL TEMED (1 μL/mL of gel), followed by APS 5 μL (40% stock concentration). The solution was mixed thoroughly and poured between glass plates with 0.75 mm spacers, followed by insertion of a 0.75 mm comb. The gel was then left to fully polymerise. Approximately 80% of the cassette was filled with TBE buffer. The comb was removed after polymerisation and the wells were flushed with deionised water, and then buffer. 10 μL at 20 nM of each aptamer sample in water was prepared along with 10 μL urea (8 M). 20 μL of each sample was loaded onto the polyacrylamide gel. All gels were run at 300 V, 15 mA, for 60 minutes. Gels were stained in Stains-All prepared in isopropanol-tris buffer, for over an hour. The gels were then rinsed to remove excess stain using water before being imaged on an Epsom scanner.

### 4.3.23. PAGE-based Purification of Aptamers

MinE07-Biotin, MinE07M3-Biotin, MinE07M6-Biotin, MinE07-U-Ph-Biotin, MinE07-U-Vi-Biotin and MinE07-U-I-Biotin aptamers required purification by Poly Acrylamide Gel Electrophoresis (PAGE). A 1.5 mm 15% denaturing PAGE gel was produced by diluting 37.5 mL of 20% acrylamide denaturing stock solution with 12.5 mL of 1 x TBE buffer. To this 50 μL of TEMED and 130 μL of

40% APS stock solution were added to induce polymerisation. The gel was poured between two glass plates and a 1.5 mm comb was inserted. The comb was removed after casting and the well washed with TBE buffer. The gel was pre-run at 300 V for 1 hour before the sample was loaded. The aptamers with 8 M urea were loaded onto the gel and run at 250 V for half an hour before being run for a further 1.5 hours at 300 V. The gel was removed from the glass plates and placed onto cling film, which was then placed onto a silica TLC plate; the gel was illuminated with UV light to visualise each aptamer. The aptamer band was cut out of the bulk gel and placed into a 15 mL falcon tube. The gel was homogenised, and 10 mL of autoclaved water was added. The solution was mixed and rapidly frozen in liquid nitrogen before being incubated overnight at 60 °C in a water bath. The solution was split into two equal portions and centrifuged for five minutes, the supernatants were collected, and the pellet extracted using a pipette with 2 mL autoclaved water. This was repeated three times. The supernatants were dried by centrifugal vacuum concentration at 60 °C for five hours and the pellets were re-suspended in autoclaved water.

### 4.3.24. Ethanol Precipitation of Aptamers after PAGE Purification

MinE07, MinE07M3 and MinE07M6 aptamers were desalted by ethanol precipitation, the samples were incubated with 3 M sodium acetate and 100% ethanol over night before being centrifuged for 20 minutes. The supernatant was removed, and the pellet washed with 1 mL of ethanol and centrifuged again; this was repeated three times for each sample. The samples were then left to air dry before being re-suspended in autoclaved water.

### 4.3.25. Determination of Aptamer concentration by UV-Visible spectrophotometry

Aptamer concentrations were analysed by UV-Visible spectrometry using a Nanodrop spectrophotometer. 2 µL of deionised water was placed onto the stage and this blank was run through the machine. The stage was cleaned with deionised water and a Kimtech wipe. 2 µL of sample was then placed onto the stage and a spectrum was run between the regions of 200 –

360 nm. Each sample was repeated until a minimum of three concordant results were achieved.

Using the calculated $A_{260}$ from Intregrated DNA Technologies (IDT) (2.11) and the measured $A_{260}$ from the UV-Vis data, a concentration was calculated by multiplying them together.

# 4.4. Results and Discussion

## 4.4.1. Synthesis of MinE07 aptamer

### 4.4.1.1. Synthesis of MinE07 aptamer

MinE07: 5'-rGrGrAfCrGrGrAfUfUfUrArAfUfCrGfCfCrGfUrArGrArArArArGfCrAfUrGfUfCrArArArG fCfCrGrGrArAfCfCrGfUfCfC-3'

The synthesis of **MinE07** was done using a Expedite™ 8909 DNA synthesiser, with industrially made phosphoramidites in acetonitrile. A Universal UnyLinker solid support was used (figure 4.10).
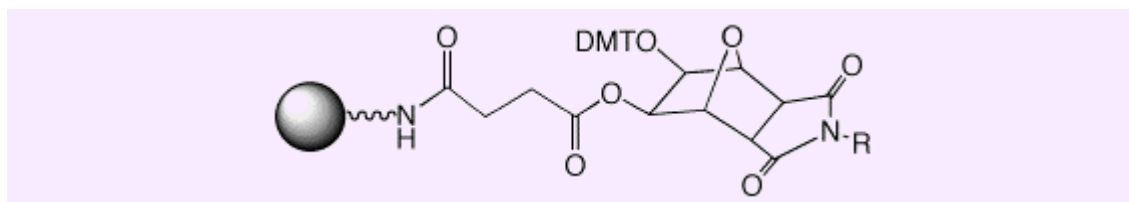


Figure 4.10. Structure of Universal UnyLinker Solid Support.

The synthesis of this aptamer took several attempts with multiple protocol changes to achieve a successful synthesis, including extended coupling times and increased washes between synthesis steps (figure 4.11).

```
 Cycle A (Adenosine) RNA 1 umole
/* ---------------------------------------------------------------------------- */
/*        Function              Mode  Amount Time(sec)        Description        */
/*                                    /Arg1   /Arg2                              */
/* ---------------------------------------------------------------------------- */
$Deblocking
 144 /*Index Fract. Coll. */ NA       1      0  "Event out ON"
   0 /*Default            */ WAIT     0     1.5 "Wait"
 141 /*Trityl Mon. On/Off */ NA       1      1  "START data collection"
  16 /*Dblk               */ PULSE   10      0  "Dblk to column"
  16 /*Dblk               */ PULSE   50     60  "Deblock"
  38 /*Diverted Wsh A     */ PULSE   40      0  "Flush system with Wsh A"
 141 /*Trityl Mon. On/Off */ NA       0      1  "STOP data collection"
  38 /*Diverted Wsh A     */ PULSE   40      0  "Flush system with Wsh A"
 144 /*Index Fract. Coll. */ NA       2      0  "Event out OFF"
$Coupling
   1 /*Wsh                */ PULSE    5      0  "Flush system with Wsh"
   2 /*Act                */ PULSE    5      0  "Flush system with Act"
  18 /*A + Act            */ PULSE    6      0  "Monomer + Act to column"
  18 /*A + Act            */ PULSE    9    402  "Couple monomer"
   1 /*Wsh                */ PULSE    8    357  "Couple monomer"
   1 /*Wsh                */ PULSE    7      0  "Flush system with Wsh"
$Capping
  12 /*Wsh A              */ PULSE   20      0  "Flush system with Wsh A"
  13 /*Caps               */ PULSE    7      0  "Caps to column"
  13 /*Caps               */ PULSE    6     15  "Cap"
  12 /*Wsh A              */ PULSE    6     15  "Cap"
  12 /*Wsh A              */ PULSE   14      0  "Flush system with Wsh A"
$Oxidizing
  15 /*Ox                 */ PULSE   20      0  "Ox to column"
  12 /*Wsh A              */ PULSE   15      0  "Flush system with Wsh A"
$Capping
  13 /*Caps               */ PULSE    7      0  "Caps to column"
  12 /*Wsh A              */ PULSE   30      0  "End of cycle wash"
```

Figure 4.11. Synthesiser Protocol for Universal UnyLinker Solid Support.

The trityl monitor of the successful synthesis of **MinE07** is shown in figure 4.12. This synthesis was used as the method development for all other aptamer synthesis using the Universal Unylinker Solid Support.
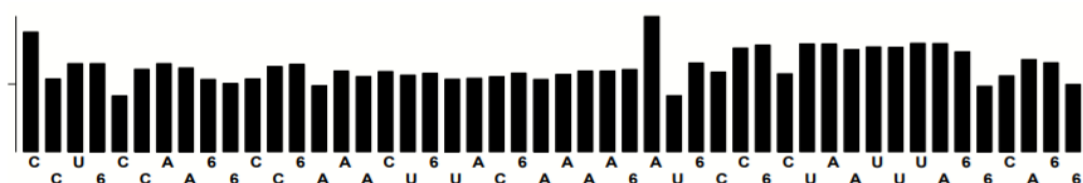


Figure 4.12. Trityl Monitor of the Synthesis of MinE07 on Universal UnyLinker Solid Support. Guanine is labelled as 6.

The step wise coupling efficiency is 97.3 as shown in the synthesis report (figure 4.13). To get the overall average yield: *$0.979^{48} = 0.361$.* This gives a 36.1% overall synthetic yield. This yield could be better however is high enough to move to carry this product forward.

Figure 4.13. Aptamer Synthesis Report of the Synthesis of MinE07 on Universal UnyLinker Solid Support.

### 4.4.1.2. 2D Predicted secondary structure of MinE07 aptamer

ViennaRNA was used to model a 2D predicted secondary structure of the **MinE07** aptamer (rna.tbi.univie.ac.at/). This was done by using the thermodynamic structure prediction and the RNAfold web server. The results for the optimal secondary structure gave a minimum free energy of -9.30kcal/mol (figure 4.14).
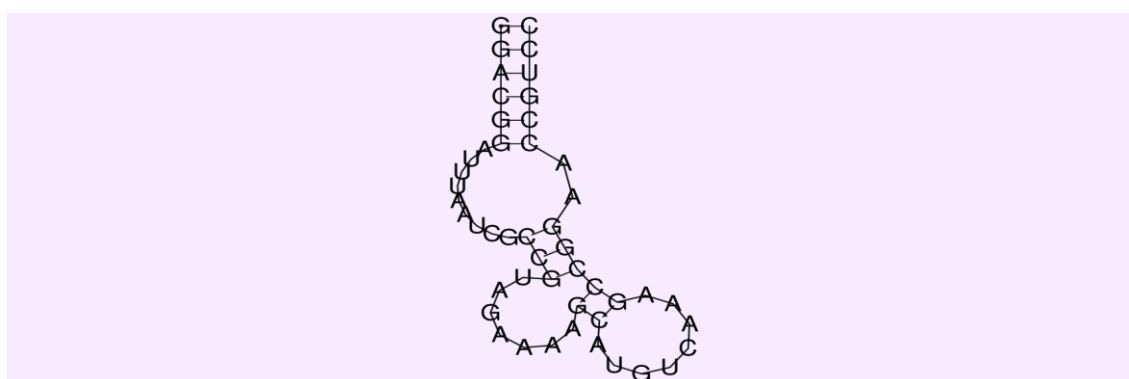


Figure 4.14. 2D Predicted secondary structure of MinE07 aptamer.

### 4.4.1.3. 3D Predicted secondary structure of MinE07 aptamer

The dot and bracket sequence was used to generate a PDB file to create 3D structure using RNA composer (rnacomposer.cs.put.poznan.pl/).

((((((.........(((.......((........)))))..))))))

As MinE07 has fluorine at the 2' position of all the uridines and cytidines residues, the PDB file was edited using notepad. All the 2'OH atoms are deleted from uridine and cytidines (figure 4.15).

```
ATOM    285  C2'   U A   9        -2.379 -17.851 -16.606  1.00  0.00           C
ATOM    286  H2'   U A   9        -2.042 -16.833 -16.402  1.00  0.00           H
ATOM    287  O2'   U A   9        -1.492 -18.832 -16.101  1.00  0.00           O
ATOM    288  HO2'  U A   9        -1.937 -19.252 -15.364  1.00  0.00           H
```

Figure 4.15. PBD File of MinE07 aptamer before editing.

The O2' lines were then replaced with F2' (figure 4.16).

```
ATOM    283  C2' U   A   9        -2.332 -17.991 -16.814  1.00  0.00           C
ATOM    284  H2' U   A   9        -1.836 -17.068 -16.520  1.00  0.00           H
ATOM    285  F2' U   A   9        -1.513 -19.038 -16.468  1.00  0.00           F
```

Figure 4.16. PBD File of MinE07 aptamer after editing.

This PDB can be visualised using Avogadro (avogadro.cc). Because the PDB file has been edited with modified structures, the geometry needs to be optimised. This was done using the optimise geometry option in Avogadro with force field MMFF94 (which assumes a vacuum) and then visualised using Hermes (figure 4.17).
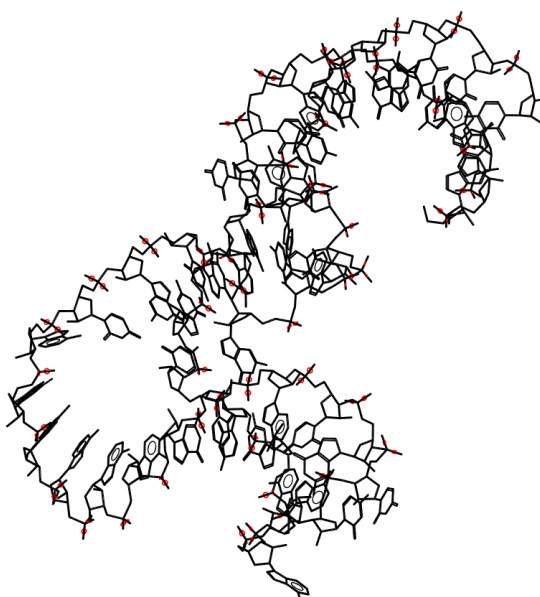


Figure 4.17. 3D Predicted secondary structure of MinE07 aptamer using Hermes from CCDC.

## 4.4.2. Synthesis of Control Aptamers MinE07- Biotin, MinE07M6-Biotin and MinE07M3-Biotin.

### 4.4.2.1. Synthesis of MinE07- Biotin aptamer

MinE07-Biotin: 5'-Biotin-rGrGrAfCrGrGrAfUfUfUrArAfUfCrGfCfCrGfUrArGrArArArArGfCrAfUrGf UfCrArArArGfCfCrGrGrArAfCfCrGfUfCfC-3'

After the method development for the synthesis of **MinE07** derivatives, a biotin tag (figure 4.18) was attached at the 5' end of the sequence so that these aptamers could be used in protein binding assays to bind to streptavidin.
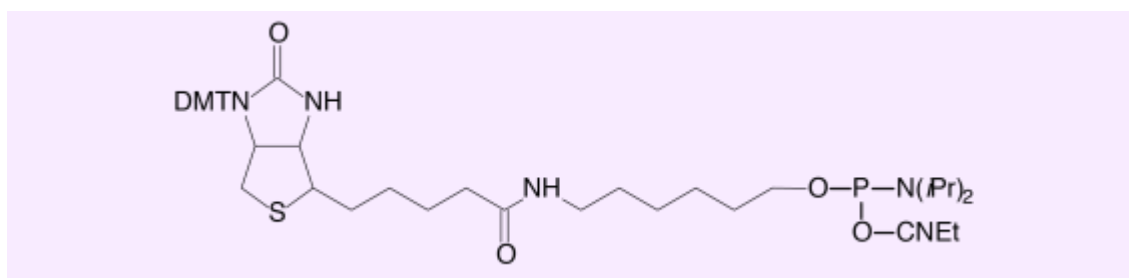


Figure 4.18. Structure of Biotin Phosphoramidite.

An original parent MinE07 aptamer was synthesised so that the quality of synthesis using the Expedite™ DNA synthesiser could be compared with an industrially made **MinE07**. This synthesis was successful, as indicated by the trityl monitor graph (figure 4.19).
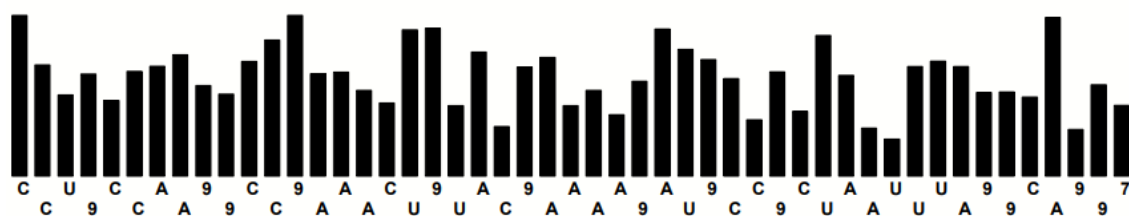


Figure 4.19. Trityl Monitor of the Synthesis of Biotin-MinE07 on Universal UnyLinker Solid Support. Guanine is labelled as 9 and the biotin is labelled as 7.

The step-wise coupling efficiency was 97.3% (figure 4.20). To get the overall average yield:

*0.973[49] = 0.261.* This gives a 26.1% overall synthetic yield.

**Sequence {5'-3'}**

799 AC9 9AU UUA AUC 9CC 9UA 9AA AA9 CAU 9UC AAA 9CC 99A ACC
9UC C

**Trityl Histogram for synthesis: AlixRNAaptamerBIOTIN9  49 mer.**
Efficiency(Stepwise Yield by base leveling): 97.3%
Final Yield (by rolling average): 100.0%
Stepwise Yield for Monomer: C:  86.3 %
Stepwise Yield for Monomer: U:  100.0 %
Stepwise Yield for Monomer: 9:  99.9 %
Stepwise Yield for Monomer: A:  100.0 %
Stepwise Yield for Monomer: 7:  100.0 %

Figure 4.20. Aptamer Synthesis Report of the Synthesis of MinE07-Biotin on Universal UnyLinker Solid Support.

### 4.4.2.2. Synthesis of MinE07M3- Biotin aptamer

MinE07M3-Biotin: 5'-Biotin-rGrGrAfCrGrGrAfUfUfUrArAfUfCrGfCfCrGfUrArGrArArArArGfCrAfU

rGfUfCrArArArGfCfCrGrGrArAfCfCrG-3'

 **MinE07M3** aptamer is 3 nucleotides shorter than the parent **MinE07** aptamer, being truncated

at the 3' end. This aptamer was synthesised because it is known to still bind to EGFR, however

the binding is weaker than the parent **MinE07** aptamer. This **MinE07M3** aptamer was used as

a control aptamer in the protein binding assays as the behaviour of it is known and comparable.

This synthesis was successful, as shown by the trityl monitor for this aptamer (figure 4.21).

C A 6 C 6 A C 6 A 6 A A A 6 C C A U U 6 C 6 7
6 C A 6 C A A U U C A A 6 U C 6 U A U A 6 A 6
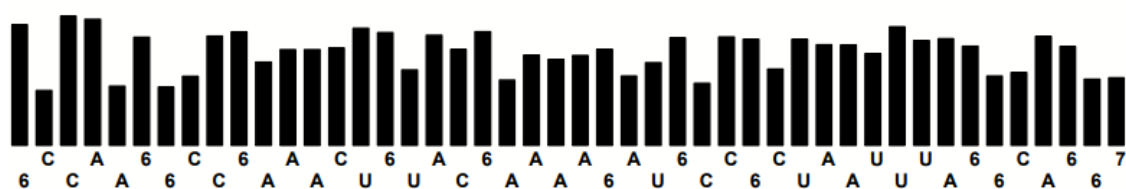
labelled as 9 and the biotin is labelled as 7.

The step-wise coupling efficiency is 99.9% as shown in the synthesis report (figure 4.22). To get the overall average yield: *0.999⁴⁶ = 0.955.* This gives a 95.5% overall synthetic yield, which is a great result.

**Sequence {5'-3'}**
766 AC6 6AU UUA AUC 6CC 6UA 6AA AA6 CAU 6UC AAA 6CC 66A ACC
6

**Trityl Histogram for synthesis: AlixRNAM3aptamerBIOTIN  46 mer.**
Efficiency(Stepwise Yield by base leveling): 100.0%
Final Yield (by rolling average): 100.0%
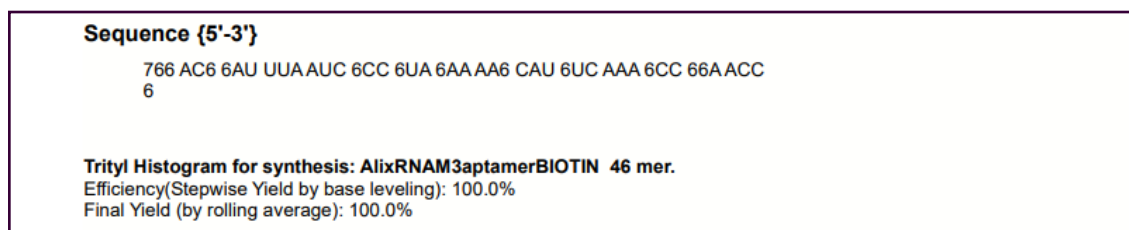
Figure 4.22. Aptamer Synthesis Report of the Synthesis of MinE07M3-Biotin on Universal UnyLinker Solid Support.

### 4.4.2.2.1. 2D/3D Predicted secondary structure of MinE07M3 aptamer

As with MinE07, ViennaRNA was used to model a 2D predicted secondary structure of the MinE07M3 aptamer (rna.tbi.univie.ac.at/). The results for the optimal secondary structure gave a minimum free energy of -3.80kcal/mol (figure 4.23).
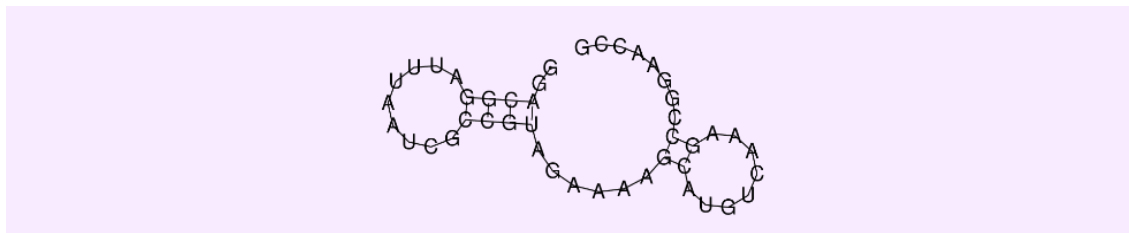


Figure 4.23. 2D Predicted secondary structure of MinE07M3 aptamer.

The dot and bracket sequence was used to generate a PDB file to create 3D structure using RNA composer (rnacomposer.cs.put.poznan.pl/).

..((((.........))))......((........))........

The same edits were made to the PDB as with **MinE07**. All the 2'OH lines were deleted and the O2' were replaced with F2' for all uridine and cytidine residues. The geometry was optimised in Avogadro with force field MMFF94 in vacuum and visualised using Hermes CCDC (figures 4.24 and 4.25).
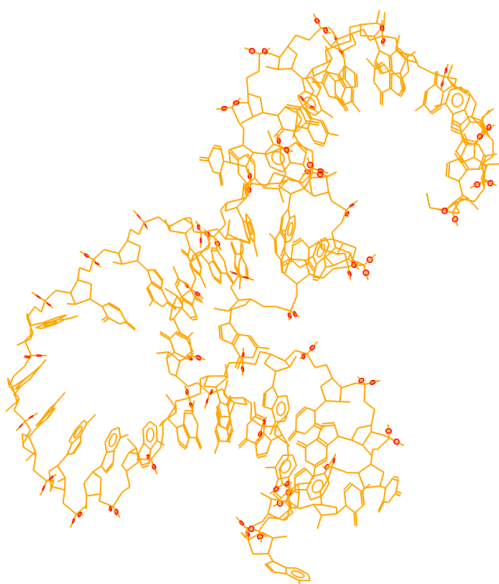
Figure 4.24. 3D Predicted secondary structure of MinE07M3 aptamer.
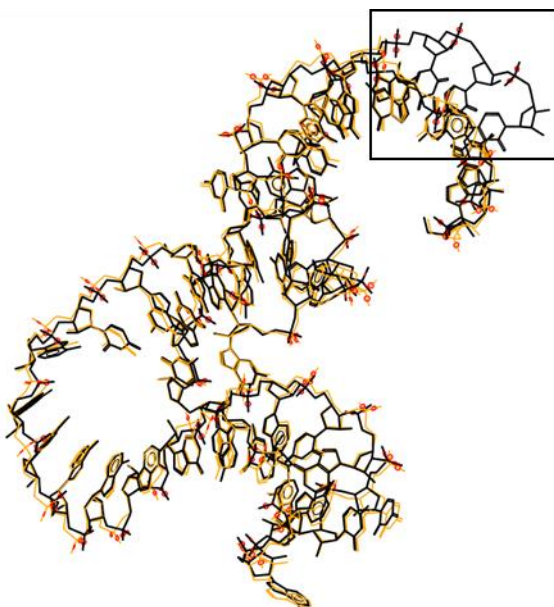


Figure 4.25. 3D Predicted secondary structure of MinE07M3 aptamer (Orange) overlaid with MinE07 (Black). The black box indicates the difference between the two structures

### 4.4.2.3. Synthesis of MinE07M6-Biotin aptamer

MinE07M6-Biotin: 5'-Biotin-rGrGrAfCrGrGrAfUfUfUrArAfUfCrGfCfCrGfUrArGrArArArArGfCrAfU rGfUfCrArArArArGfCfCrGrGrArA-3'

**MinE07M6** aptamer is 6 nucleotides shorter than the parent **MinE07** aptamer from the 3' end. This aptamer was synthesised because it is known to still bind to EGFR, however the binding is weaker than the parent **MinE07** aptamer and MinE07M3. This **MinE07M6** aptamer was used as a control aptamer, as with **MinE07M3**. This synthesis was successfully as shown by the trityl monitor for this aptamer (figure 4.26).
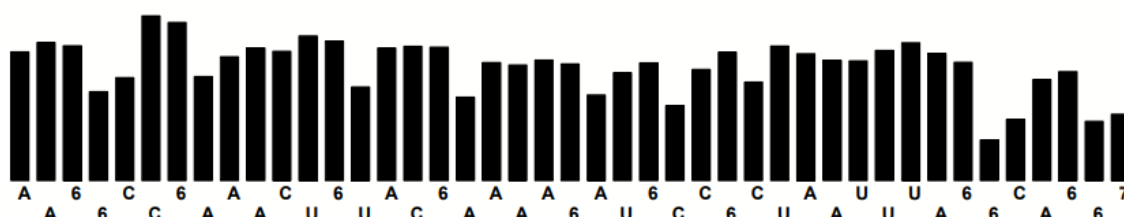


Figure 4.26. Trityl Monitor of the Synthesis of MinE07M6-Biotin on Universal UnyLinker Solid Support. Guanine is labelled as 6 and the biotin is labelled as 7.

The step wise coupling efficiency is 98.65% as shown in the synthesis report (figure 4.27). To get the overall average yield: *0.986[43] = 0.545*. This gives a 54.5% overall synthetic yield, which is a good enough result to be carried forward.



**Sequence {5'-3'}**

766 AC6 6AU UUA AUC 6CC 6UA 6AA AA6 CAU 6UC AAA 6CC 66A A

**Trityl Histogram for synthesis: AlixRNAM6aptamerBIOTIN 43 mer.**
Efficiency(Stepwise Yield by base leveling): 98.6%
Final Yield (by rolling average): 100.0%
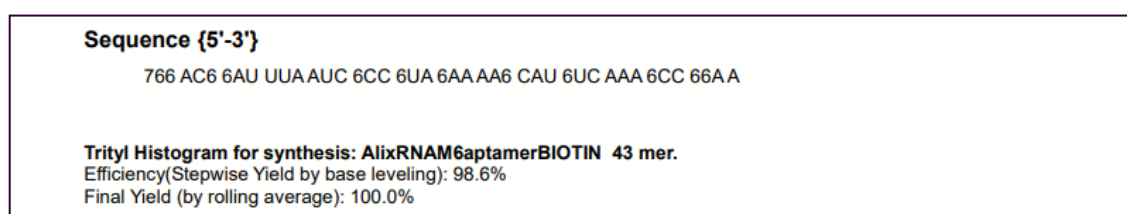
Figure 4.27. Aptamer Synthesis Report of the Synthesis of MinE07M6-Biotin on Universal UnyLinker Solid Support.

#### 4.4.2.3.1. 2D/3D Predicted secondary structure of MinE07M6 aptamer

As with MinE07 and MinE07M3, ViennaRNA was used to model a 2D predicted secondary structure of MinE07M6 aptamer (rna.tbi.univie.ac.at/). The results for the optimal secondary structure gave a minimum free energy of -3.80kcal/mol also (figure 4.28).
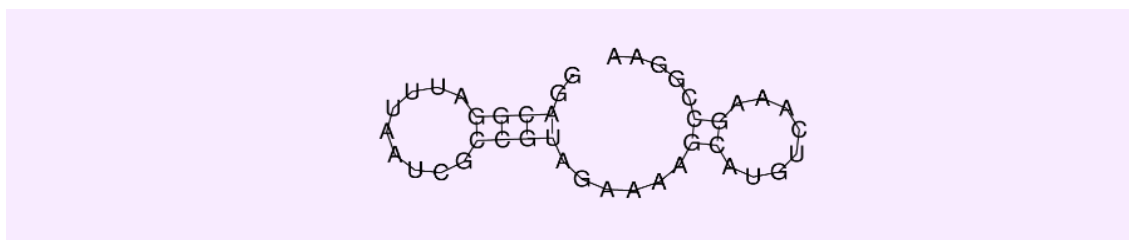
Figure 4.28. 2D Predicted secondary structure of MinE07M6 aptamer.

The dot and bracket sequence was used to generate a PDB file to create 3D structure using RNA composer (rnacomposer.cs.put.poznan.pl/).

..((((.........))))......((........))....

The same edits and optimisation steps were done as with MinE07 and MinE07M3 (figures 4.29 and 4.30).
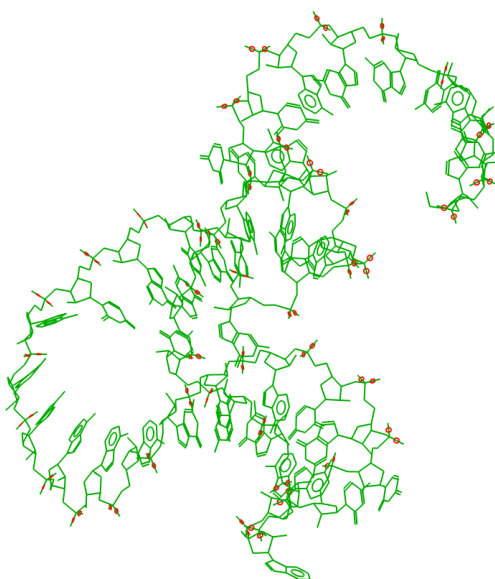


Figure 4.29. 3D Predicted secondary structure of MinE07M6 aptamer.
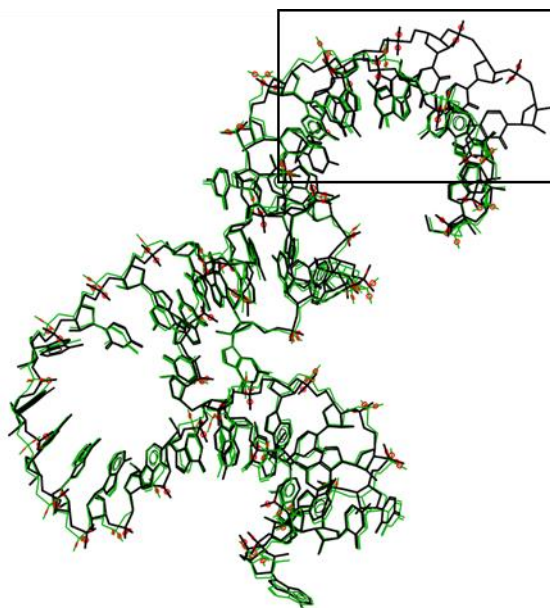
Figure 4.30. 3D Predicted secondary structure of MinE07M6 aptamer (Green) overlaid with MinE07 (Black). The black box indicates the difference between the two structures.

This sample was then cleaved from the solid support and purified by PAGE purification. The sample was then desalted with Zetadex resin and run on an analytical polyacrylamide gel to check that the molecular weight was correct for the length of this aptamer.

4.4.2.4. Purification of MinE07-Biotin, MinE07M3-Biotin and MinE07M6-Biotin

All of these aptamer samples were then cleaved from the solid support and purified by PAGE. The samples were then desalted with Zetadex resin and run on an analytical polyacrylamide gel to check that the molecular weight was correct for the length of this aptamer.

4.4.2.4.1. Analytical Polyacrylamide Gel

To check the success of the synthesis of **MinE07-Biotin, MinE07M3-Biotin** and **MinE07M6-Biotin** a 15% polyacrylamide gel was run after cleavage from the solid support and before purification of the samples (figure 4.31).
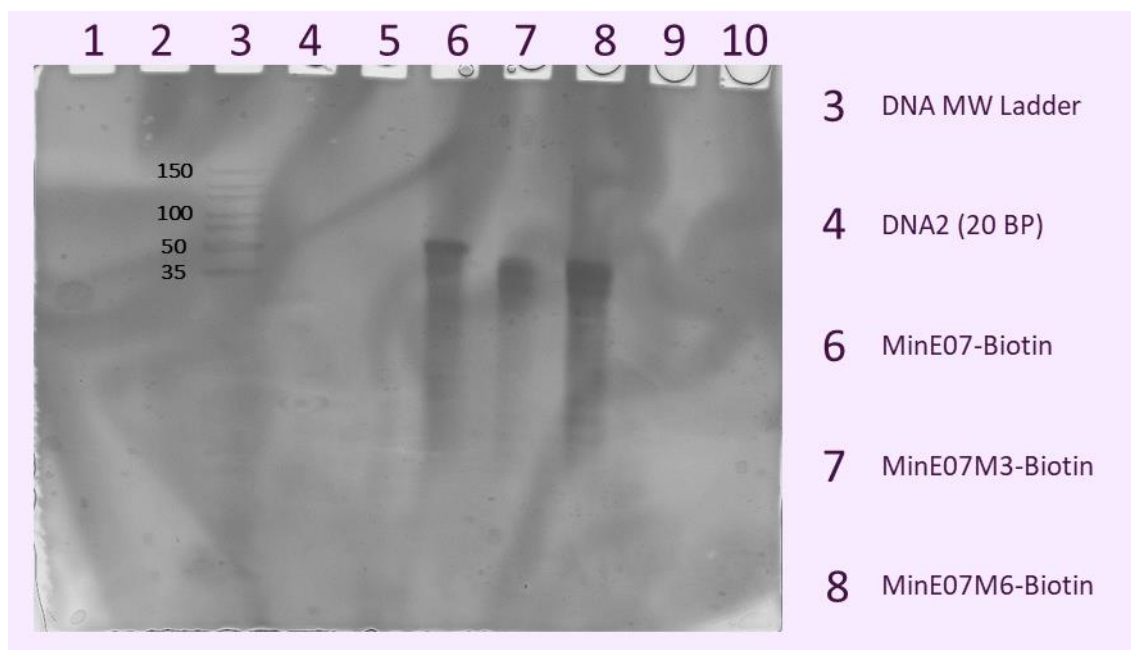
Figure 4.31. 15% Polyacrylamide Gel of MinE07-Biotin, MinE07M3-Biotin and MinE07M6-Biotin before PAGE purification .

As shown by figure 4.31, **MinE07-Biotin, MinE07M3-Biotin** and **MinE07M6-Biotin** have their top bands showing the correct molecular weight. However, there is substantial amount of non-full-length sequences also present, as indicated by the smear below each top band. These samples were then purified by PAGE purification, cutting out the correct length bands to get a single sequence sample. These purified samples were then desalted with Zetadex resin and then run again on an analytical polyacrylamide gel to check that the PAGE purification was successful (figure 4.32).
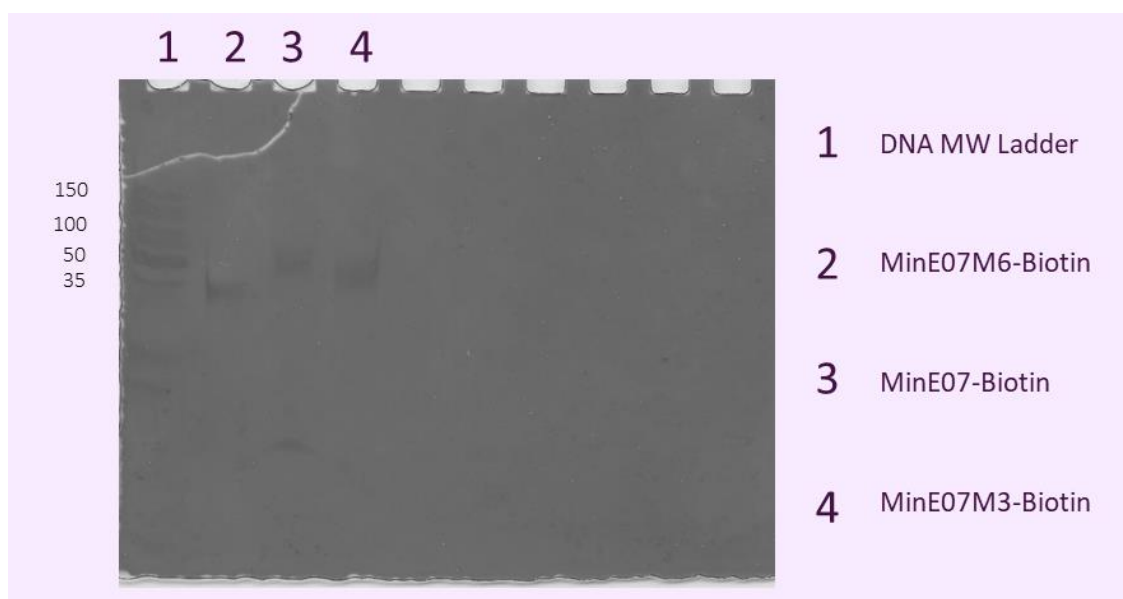
Figure 4.32. 15% Polyacrylamide Gel of MinE07-Biotin, MinE07M3-Biotin and MinE07M6-Biotin after PAGE

purification.

As shown by figure 4.32, the PAGE purification was potentially successful, however to confirm

the structure a HABA assay will be run on the aptamers to check their was a biotin on the end

(see chapter 5).

### 4.4.3. Synthesis of MinE07 Library

MinE07Lib: 5'-rGrGrAfCrGrGrAfU(U-Ph,U-Vi,U,U-I)fU(U-Ph,U-Vi,U,U-I)fU(U-Ph,U-Vi,U,U-I)rArAf

U(U-Ph,U-Vi,U,U-I)fCrGfCfCrGfU(U-Ph,U-Vi,U,U-I)rArGrArArArArGfCrAfU(U-Ph,U-Vi,U,U-I)rGfU(

U-Ph,U-Vi,U,U-I)fCrArArArGfCfCrGrGrArAfCfCrGfU(U-Ph,U-Vi,U,U-I)fCfC-PCLinker-3'
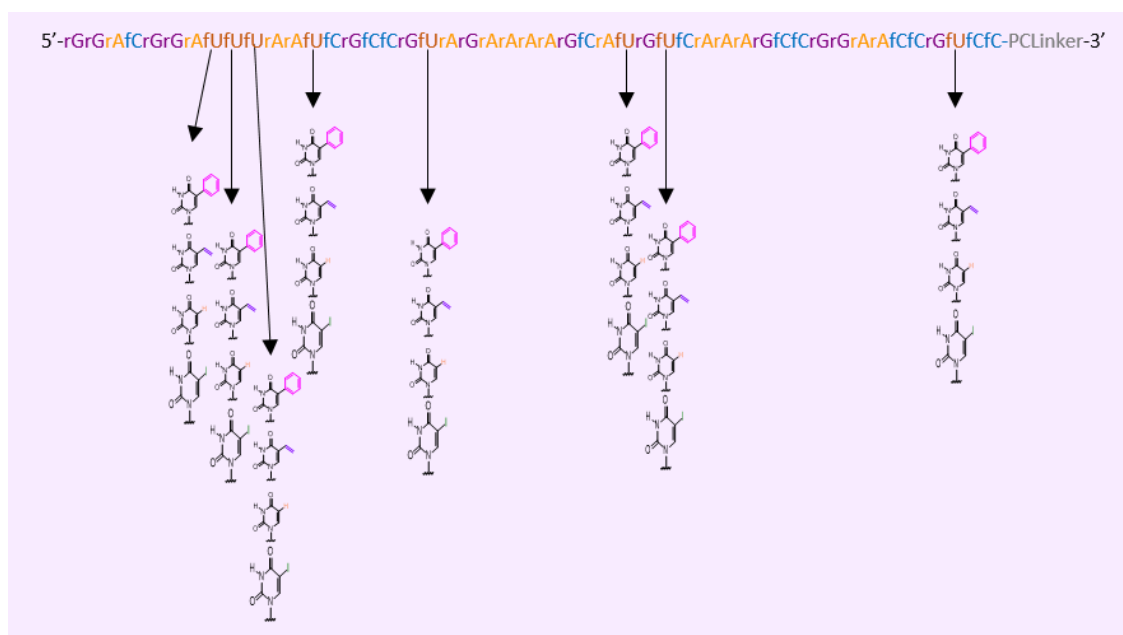


Figure 4.33. Illustration of MinE07Lib.

The MinE07Lib synthesis was conducted using an Expedite™ 8909 DNA synthesiser with commercially made phosphoramidites and modified uridines compounds **3.21**, **3.22**, **3.22**, (see chapter 3 for nucleotide synthesis). -OH modified TentaGel microspheres were used as the solid support (see chapter 2 for synthesis). Enough -OH modified TentaGel microspheres were used to create 200 copies of every possible sequence to allow for loss of microspheres during the split and mix process and synthesis errors that occur creating shorter length chains than desired. This method will make $4^8 = 65,536$ unique aptamer sequences. These microspheres were added to a synthesiser column and flushed with acetonitrile to allow them to swell (see experimental method 4.3.1 for the setup of the synthesiser).

### 4.4.3.2. Photo-cleavable Linker
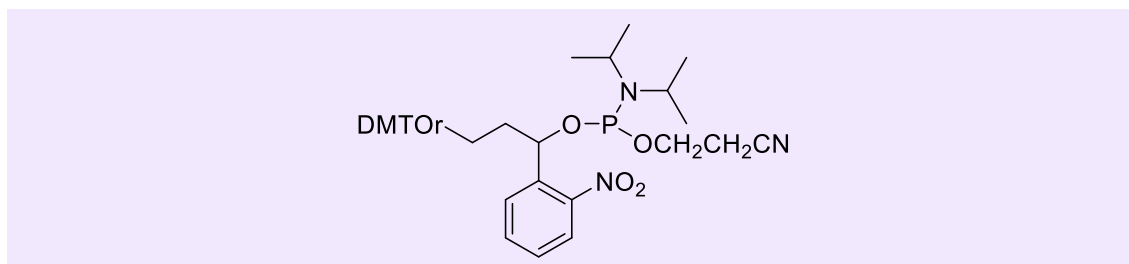


Figure 4.34. Chemical Structure of the Photocleavable Linker (3-(4,4'-Dimethoxytrityl)-1-(2-nitrophenyl)-propane-1,3-diol-[2-cyanoethyl-(N,N-diiopropyl)]-phosphoramidite).

The first monomer coupled to the microspheres was the photocleavable linker (3-(4,4'-dimethoxytrityl)-1-(2-nitrophenyl)-propane-1,3-diol-[2-cyanoethyl-(N,N-diiopropyl)]-phosphoramidite) (figure 4.34). This is added first so that after the top aptamers have been sorted using the FACS, the aptamer strands can be removed from the microsphere to be analysed using mass spectrometry. This photo-cleavable linker can be cleaved using long wave UV light in the range of 340-365 nm[459,460,461,462, 463]. In figure 4.35 we can see how they chemically will be bound to the microspheres and how they will be cleaved after library sorting.
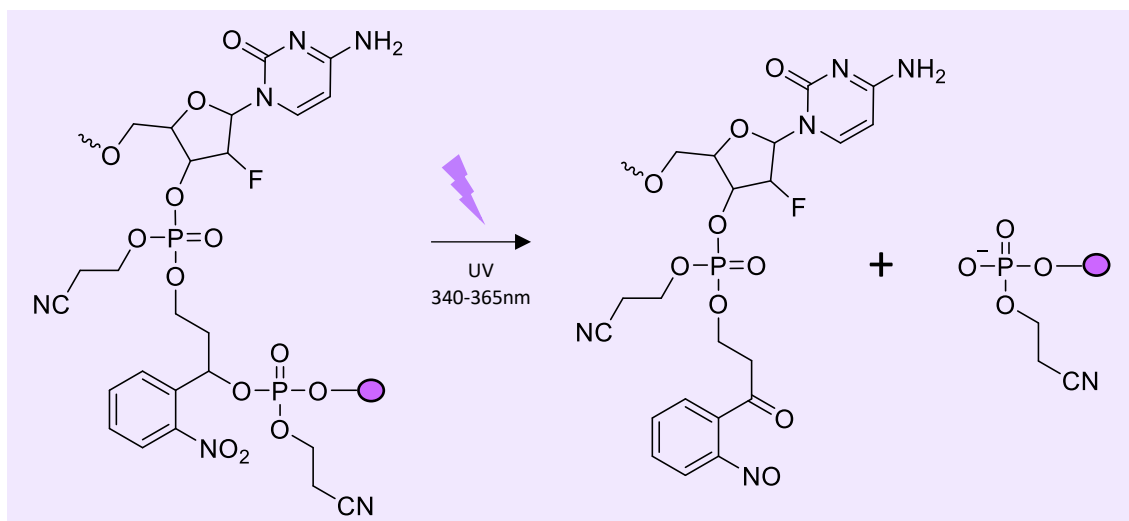


Figure 4.35. Photocleaving the microspheres from the aptamers.

### 4.4.3.3. Synthesis of MinE07Lib

All the solid support TentaGel microspheres start in the same column. After the first round of synthesis the microspheres are taken out of the main column and split 4 ways into columns U-Ph, U-Vi, U, U-I. A synthesis cycle using compound **3.21** was run through column U-Ph. The 4, 4'-dimethoxytrityl is left on fU-Ph. A synthesis cycle using compound **3.22** was run through column U-Vi. The 4, 4'-dimethoxytrityl is left on fU-Vi. A synthesis cycle using 2'-Deoxy-5'-O-DMT-2'-fluorouridine 3'-CE phosphoramidite was run through column U. The 4, 4'-dimethoxytrityl is left on fU. A synthesis circle using compound **3.23** was run through column U-I. The 4, 4'-dimethoxytrityl is left on fU-I. All the response values qualitatively indicating the coupling efficiencies for each round are displayed in table 4.1. The microspheres are taken out of column U-Ph, U-Vi, U, U-I, they are mixed together and added back into the main column. This method was repeated for split and mix round.

Table 4.1 displays the response values from the trityl monitors of each round of synthesis. It is important in the aptamer synthesis that the coupling efficiency stays as consistent as possible and $<10^6$ being the optimum value. As seen in the table some values do drop to $10^5$ suggesting that not every coupling step worked as well as others. In future work, optimising these coupling steps would improve this method by make the next selection and identification steps easier.

Table 4.1. Displaying the response values from the trityl monitors of each round of synthesis.

| Round | Nucleotides | Response value – qualitative indication of coupling efficiency |
|---|---|---|
| 1 | 3'-PCLinker-fCfC-5' | $9.52 \times 10^5$ |
| 2 | 3'- fU (U-Ph, U-Vi, U, U-I)-5' | $4.95 \times 10^5$, $6.04 \times 10^5$, $1.60 \times 10^6$, $8.49 \times 10^5$ |
| 3 | 3'-rGfCfCrArArGrGfCfCrGrArArAfC-5' | $1.28 \times 10^6$ |
| 4 | 3'- fU (U-Ph, U-Vi, U, U-I)-5' | $1.60 \times 10^6$, $4.60 \times 10^5$, $1.28 \times 10^6$, $1.66 \times 10^6$ |
| 5 | 3'-rG-5' | $2.52 \times 10^5$ |
| 6 | 3'- fU (U-Ph, U-Vi, U, U-I)-5' | $3.01 \times 10^5$, $1.36 \times 10^6$, $1.44 \times 10^5$, $8.95 \times 10^5$ |
| 7 | 3'-rAfCrGrArArArArGrA-5' | $1.47 \times 10^6$ |
| 8 | 3'- fU (U-Ph, U-Vi, U, U-I)-5' | $1.71 \times 10^6$, $2.65 \times 10^5$, $1.20 \times 10^6$, $1.69 \times 10^5$ |
| 9 | 3'- rGfCfCrGfC -5' | $1.38 \times 10^6$ |
| 10 | 3'- fU (U-Ph, U-Vi, U, U-I)-5' | $6.07 \times 10^5$, $2.59 \times 10^5$, $2.65 \times 10^5$, $4.29 \times 10^5$ |
| 11 | 3'- rArA -5' | $1.28 \times 10^6$ |
| 12 | 3'- fU (U-Ph, U-Vi, U, U-I)-5' | $2.02 \times 10^6$, $4.38 \times 10^5$, $2.86 \times 10^5$, $1.27 \times 10^6$ |
| 13 | 3'- fU (U-Ph, U-Vi, U, U-I)-5' | $8.95 \times 10^5$, $9.21 \times 10^5$, $1.54 \times 10^5$, $1.58 \times 10^5$ |
| 14 | 3'- fU (U-Ph, U-Vi, U, U-I)-5' | $3.29 \times 10^5$, $6.34 \times 10^5$, $1.64 \times 10^5$, $7.01 \times 10^5$ |
| 15 | 3'-rArGrGfCrArGrG-5' | $1.38 \times 10^6$ |

The final weight of the **MinE07Lib** TentaGel Microspheres was 0.0029g, meaning a loss of 67.4% from the original weight of 0.0087g. This was caused by the human error when moving the microspheres from column to column. This means there is now only approximately 66.6 copies of each novel aptamer sequence in **MinE07Lib** instead of the original 200 copies.

The **MinE07Lib** was synthesised to a high enough standard to be carried forward to the next aptamer selection stage. Possibly a larger starting quantity of microspheres would have been better and made up for the large amount of loss during the split and mix stages.

## 4.4.4. Synthesis of fully modified control aptamers MinE07-U-Ph-Biotin, MinE07-U-Vi-Biotin and MinE07-U-I-Biotin

Aptamers were synthesised to have all their uridines modified with a single modification. This was done to see whether these modifications increase or disrupt the high binding affinity between MinE07 and EGFR, and to be compared with the aptamers that have been selected from the library using FACS.

### 4.4.4.1. Synthesis of MinE07-U-Ph-Biotin Aptamer

MinE07-U-Ph-Biotin:5'-Biotin-rGrGrAfCrGrGrAfU-PhfU-PhfU-PhrArAfU-PhfCrGfCfCrGfU-PhrAr GrArArArArGfCrAfU-PhrGfU-PhfCrArArArGfCfCrGrGrArAfCfCrGfU-PhfCfC-3'.

**MinE07U-Ph-Biotin** aptamer is the same sequence as the parent **MinE07** aptamer however all the uridines are modified with modification Ph compound **3.21** (figure 4.36). Synthesis of compound **3.21** can be found in chapter 3.4.
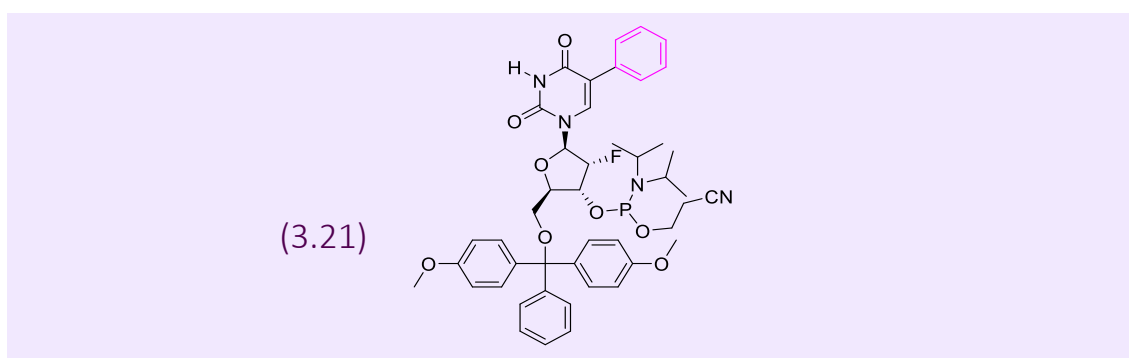


Figure 4.36. Chemical Structure of DMTr -protected 2'-desoxy-2'-fluoro-5-phenyl-uridine phosphoramidite (**3.21**).

This aptamer was synthesised to see whether modifying all the uridines influenced the binding of **MinE07** to EGFR. The binding of this aptamer with EGFR will be directly comparable with

**MinE07, MinE07M3** and **MinE07M6**. This synthesis was successfully shown by the trityl monitor

for this aptamer (figure 4.37).
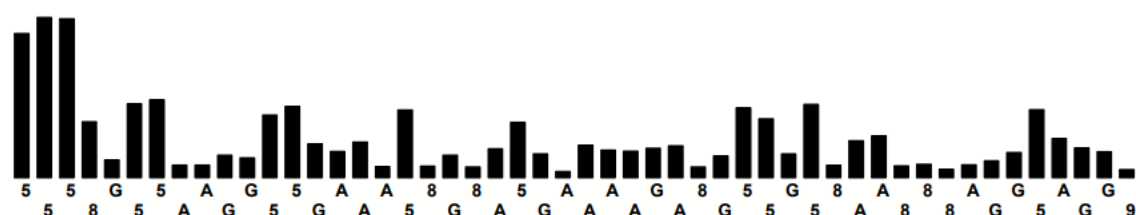


Figure 4.37. Trityl Monitor of the Synthesis of MinE07U-Ph-Biotin on Universal UnyLinker Solid Support. Cytosine is

labelled as 5, U-Ph (**3.21**) is 8 and the biotin is 9.

The step wise coupling efficiency is 98.4% as shown in the synthesis report (figure 4.38). To get

the overall average yield: *$0.984^{49} = 0.454$.* This gives a 45.4% overall synthetic yield, which is a

good enough yield to carry forward for further testing.



```
Sequence {5'-3'}
    9GG A5G GA8 88A A85 G55 G8A GAA AAG 5A8 G85 AAA G55 GGA A55
    G85 55


Trityl Histogram for synthesis: ARPBIOMINE07A  50 mer.
Efficiency(Stepwise Yield by base leveling): 98.4%
Final Yield (by rolling average): 100.0%
Stepwise Yield for Monomer: 5:  97.2 %
Stepwise Yield for Monomer: 8:  94.7 %
Stepwise Yield for Monomer: G:  100.0 %
Stepwise Yield for Monomer: A:  100.0 %
Stepwise Yield for Monomer: 9:  100.0 %
```
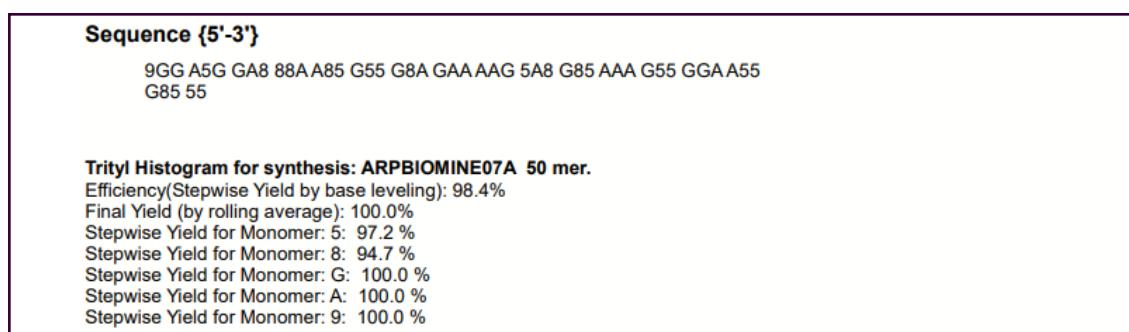
Figure 4.38. Aptamer Synthesis Report of the Synthesis of MinE07-U-Ph-Biotin on Universal UnyLinker Solid

Support.

### 4.4.4.1.1. 3D Predicted secondary structure of MinE07-U-Ph aptamer

To get the 3D predicted structure of the **MinE07-U-Ph** aptamer the same F2' edited PDB file as

MinE07 was opened in Avogardo (avogadro.cc). The hydrogen atom at C(5) on all the uridine

residues were deleted and a phenyl was added (figure 4.39).
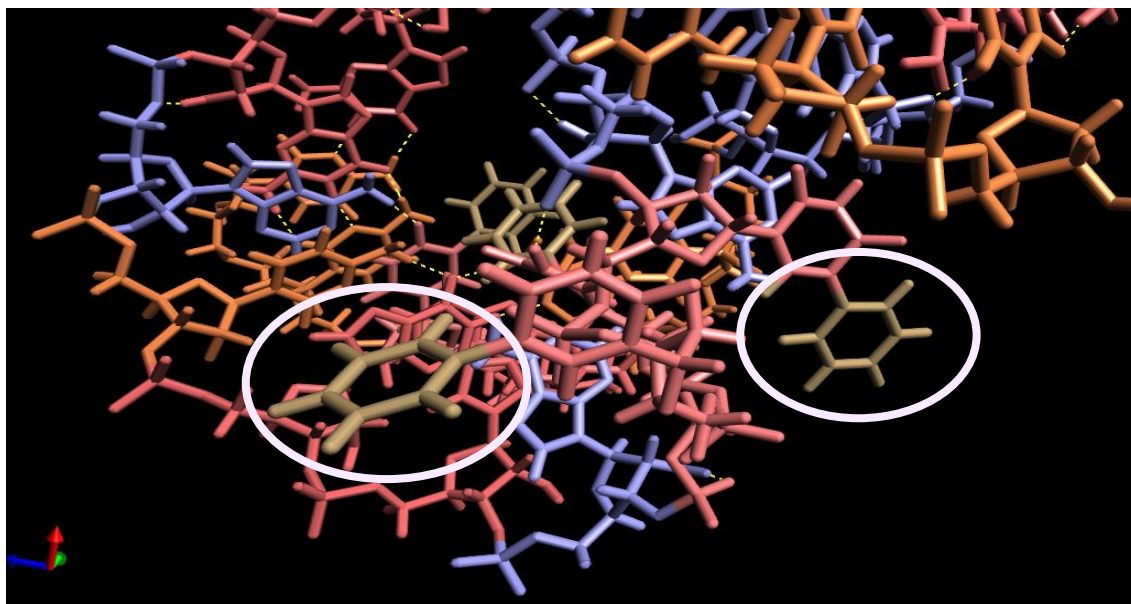
Figure 4.39. Phenyl edit Zoom in of 3D Predicted secondary structure of MinE07-U-Ph aptamer from Avogadros
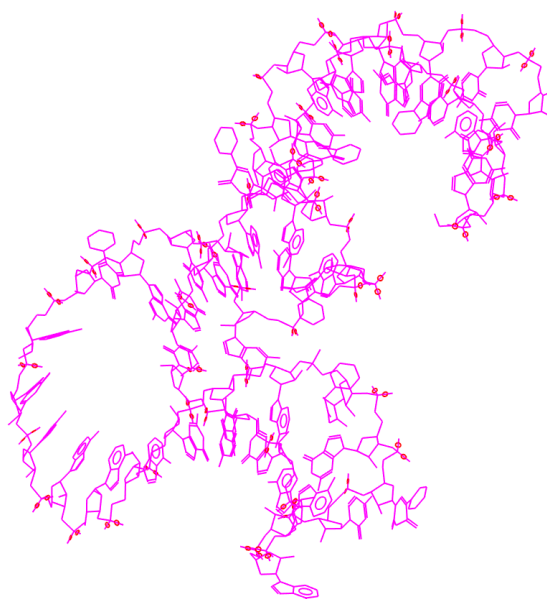


Figure 4.40. 3D Predicted secondary structure of MinE07-U-Ph aptamer in Avogadro with optimised geometry using

force field MMFF94.
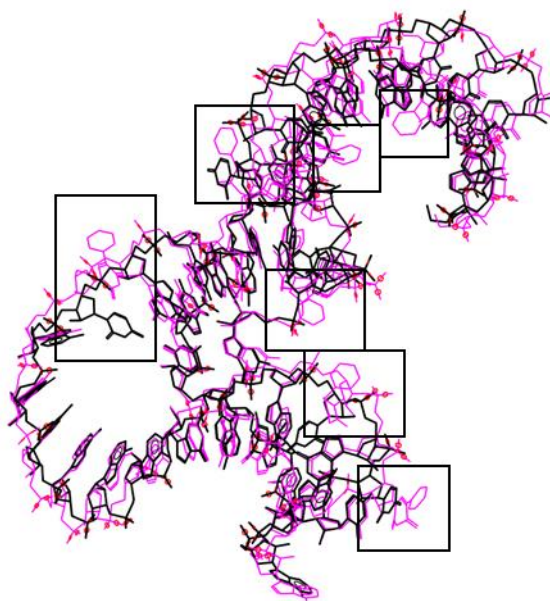
Figure 4.41. 3D Predicted secondary structure of MinE07-U-Ph aptamer (Pink) overlaid with MinE07 aptamer (Black). The black boxes displaying the difference between the two structures.

The size of the Ph on the uridines causes some of the nucleotides to rotate in different directions to fit the space. This can be visualised in figure 4.41.

4.4.4.2. Synthesis of MinE07-U-Vi-Biotin aptamer

MinE07-U-Vi-Biotin:5'-Biotin-rGrGrAfCrGrGrAfU-VifU-VifU-VirArAfU-VifCrGfCfCrGfU-VirArGrAr

ArArArGfCrAfU-VirGfU-VifCrArArArGfCfCrGrGrArArAfCfCrGfU-VifCfC-3'
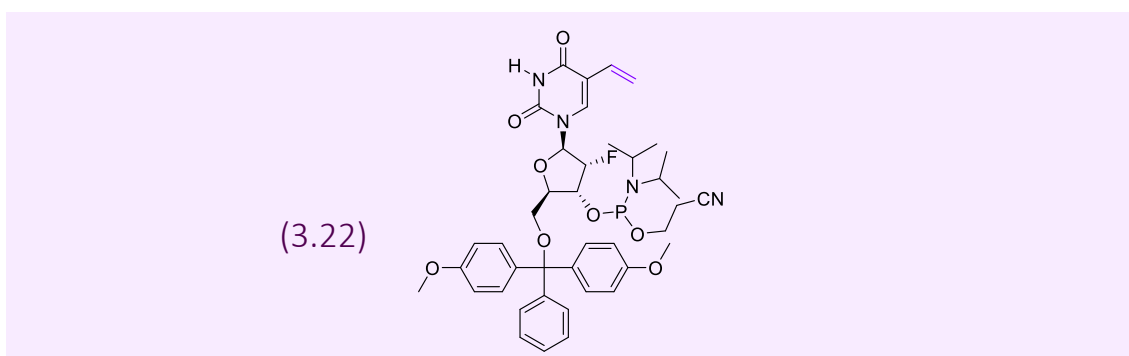


Figure 4.42. Chemical Structure of DMTr -protected 2'-desoxy-2'-fluoro-5-vinyl-uridine phosphoramidite (**3.22**).

In **MinE07-U-Vi-Biotin** aptamer, the uridines are modified with modification Vi compound **3.22** (figure 4.42). This synthesis was successfully shown by the trityl monitor for this aptamer (figure 4.43).
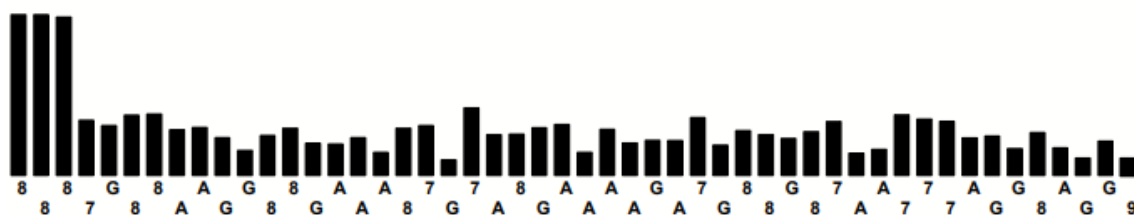
Figure 4.43. Trityl Monitor of the Synthesis of MinE07U-Vi-Biotin on Universal UnyLinker Solid Support Cytosine is labelled as 8, U-Vi (**3.22**) is 7 and biotin is 9.

The step wise coupling efficiency is 97.8% as shown in the synthesis report (figure 4.44). To get the overall average yield: *$0.978^{49}$ = 0.3362.* This gives a 33.62% overall synthetic yield. This yield could be better however it is enough to carry forward for further testing.



**Sequence {5'-3'}**

9GG A8G GA7 77A A78 G88 G7A GAA AAG 8A7 G78 AAA G88 GGA A88
G78 88

**Trityl Histogram for synthesis: ARP-BIOMINE07-U-VI  50 mer.**
Efficiency(Stepwise Yield by base leveling): 97.8%
Final Yield (by rolling average): 100.0%
Stepwise Yield for Monomer: 8:  93.6 %
Stepwise Yield for Monomer: 7:  100.0 %
Stepwise Yield for Monomer: G:  97.1 %
Stepwise Yield for Monomer: A:  98.5 %
Stepwise Yield for Monomer: 9:  100.0 %
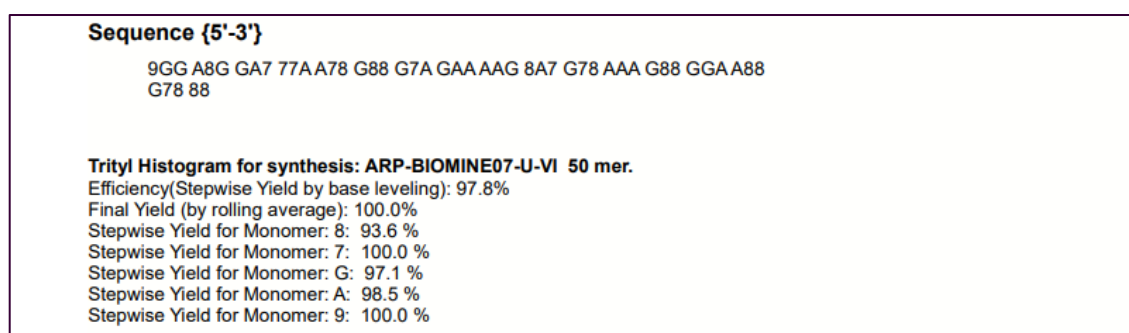
Figure 4.44. Aptamer Synthesis Report of the Synthesis of MinE07U-Vi-Biotin on Universal UnyLinker Solid Support.

### 4.4.4.2.1. 3D Predicted secondary structure of MinE07-U-Vi aptamer

To get the 3D predicted structure of **MinE07-U-Vi** aptamer the same F2' edited PDB file as MinE07 was opened in Avogadro (avogadro.cc). The hydrogen atom at C(5) on all the uridine residues were deleted and a vinyl was added (figure 4.45).

235
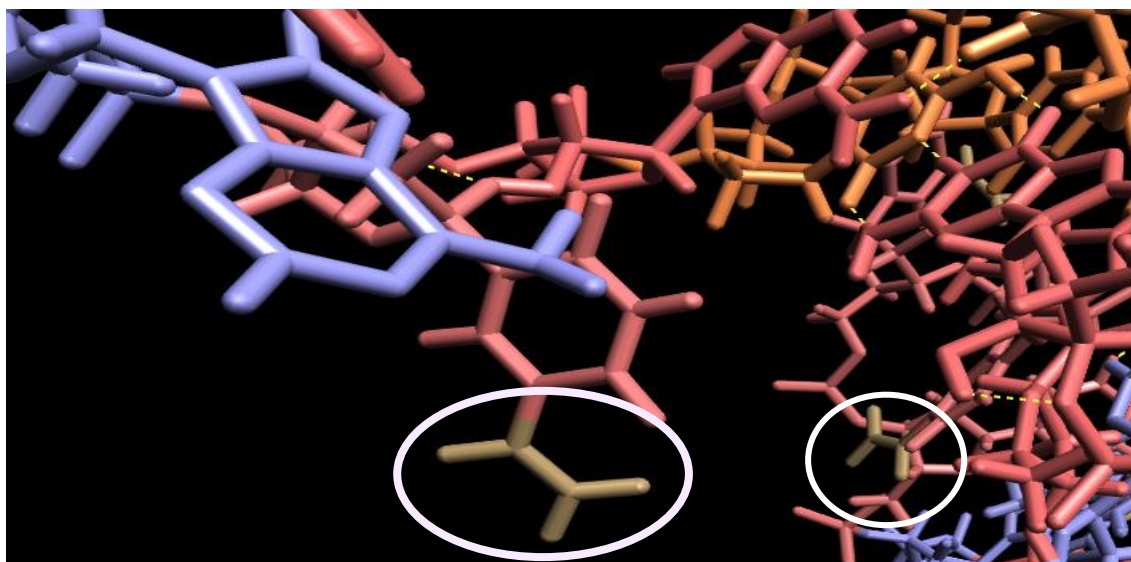
Figure 4.45. Vinyl edit Zoom in of 3D Predicted secondary structure of MinE07-U-Vi aptamer from Avogadro.
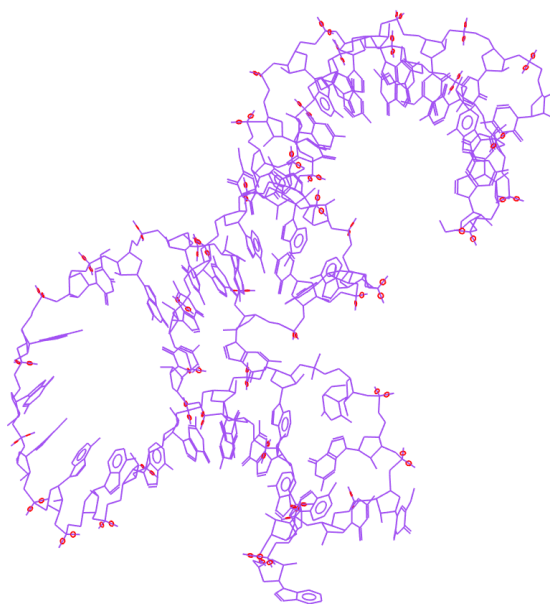


Figure 4.46. 3D Predicted secondary structure of MinE07-U-Vi aptamer in Avogadro with optimised geometry using

force field MMFF94.
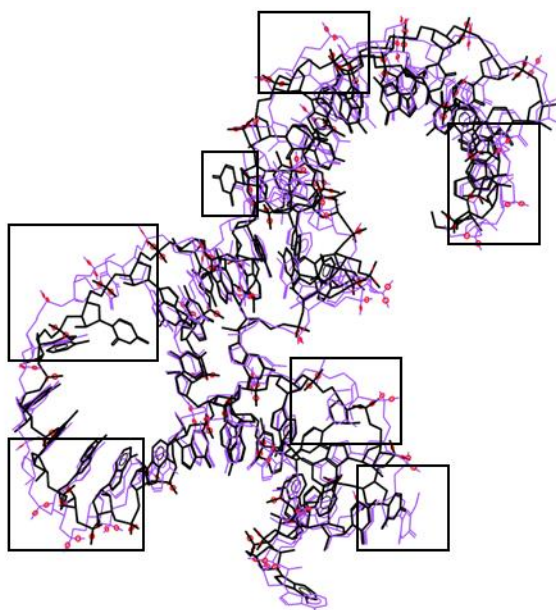
Figure 4.47. 3D Predicted secondary structure of MinE07-U-Vi aptamer (Purple) overlaid with MinE07 aptamer (Black). The black boxes displaying the difference between the two structures.

The Vi modified uridines cause the aptamer to change shape as can be seen by the difference in the purple and black aptamers in figure 4.47. The different in shape is not as large as with the Ph modification, caused by less steric hindrance.

4.4.4.3. Synthesis of MinE07-U-I-Biotin aptamer

MinE07-U-I-Biotin:5'-Biotin-rGrGrAfCrGrGrAfU-IfU-IfU-IrArAfU-IfCrGfCfCrGfU-IrArGrArArArArG fCrAfU-IrGfU-IfCrArArArGfCfCrGrGrArArAfCfCrGfU-IfCfC-3'.
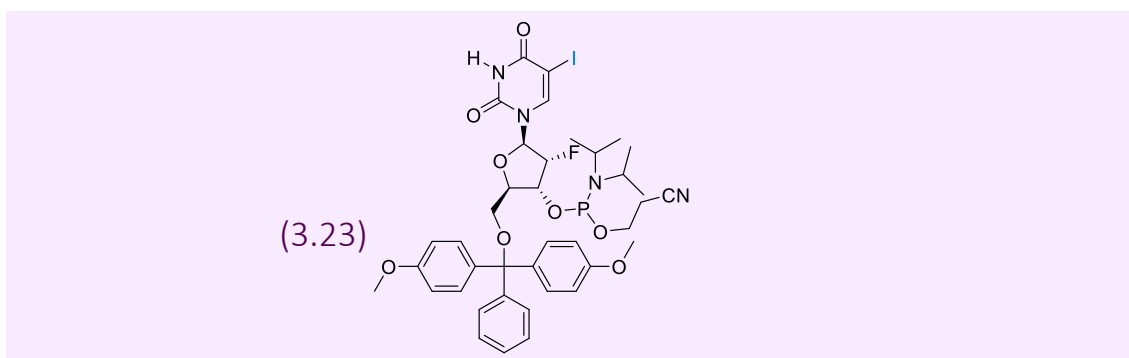


Figure 4.48. Chemical Structure of DMTr -protected 2'-desoxy-2'-fluoro-5-iodo-uridine phosphoramidite (**3.23**)

In the **MinE07-U-I-Biotin** aptamer the uridines are modified with modification I, compound **3.23** (figure 4.48). This synthesis was successfully shown by the trityl monitor (figure 4.49).

237
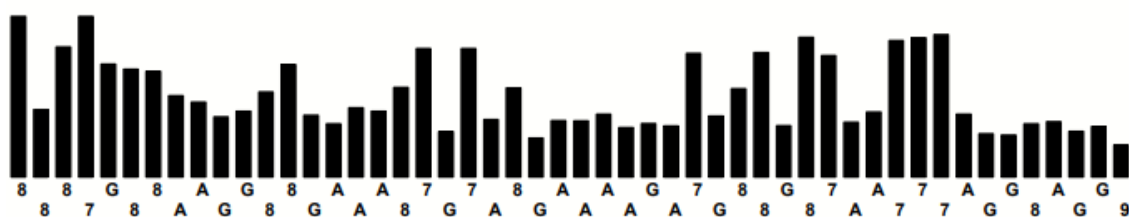
Figure 4.49. Trityl Monitor of the Synthesis of MinE07U-I-Biotin on Universal UnyLinker Solid Support Cytosine is labelled as 8, U-I (**3.23**) is 7 and biotin is 9.

The step wise coupling efficiency is 96.1% as shown in the synthesis report (figure 4.50). To get the overall average yield: *0.961$^{49}$ = 0.1424.* This gives a 14.24% overall synthetic yield. This yield is poor, however enough was produced to carry forward for further testing.

**Sequence {5'-3'}**

    9GG A8G GA7 77A A78 G88 G7A GAA AAG 8A7 G78 AAA G88 GGA A88
    G78 88

**Trityl Histogram for synthesis: ARP-BIOMINE07-U-I  50 mer.**
Efficiency(Stepwise Yield by base leveling): 96.1%
Final Yield (by rolling average): 100.0%
Stepwise Yield for Monomer: 8:  87.7 %
Stepwise Yield for Monomer: 7:  99.3 %
Stepwise Yield for Monomer: G:  95.8 %
Stepwise Yield for Monomer: A:  97.7 %
Stepwise Yield for Monomer: 9:  100.0 %
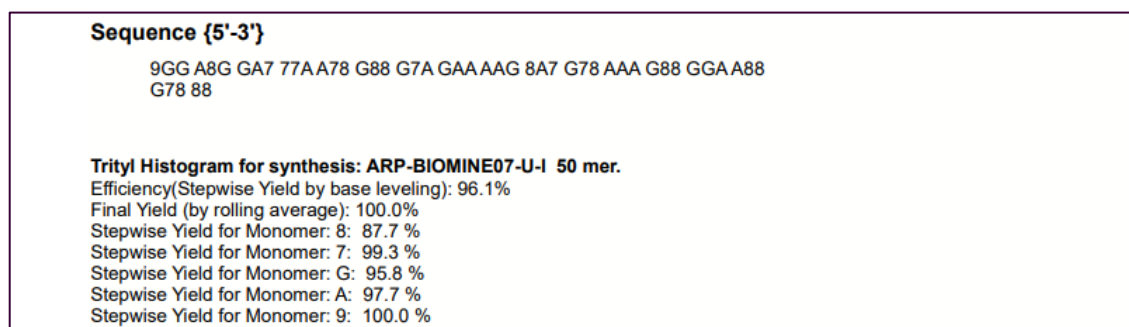
Figure 4.50. Aptamer Synthesis Report of the Synthesis of MinE07-U-I-Biotin on Universal UnyLinker Solid Support.

### 4.4.4.3.1. 3D Predicted secondary structure of MinE07-U-I aptamer

To get the 3D predicted structure of **MinE07-U-I** aptamer the same F2' edited PDB file as MinE07 was opened in Avogadro (avogadro.cc). The hydrogen atom at C(5) on all the uridine residues were deleted and an iodine atom was added (figure 4.51).
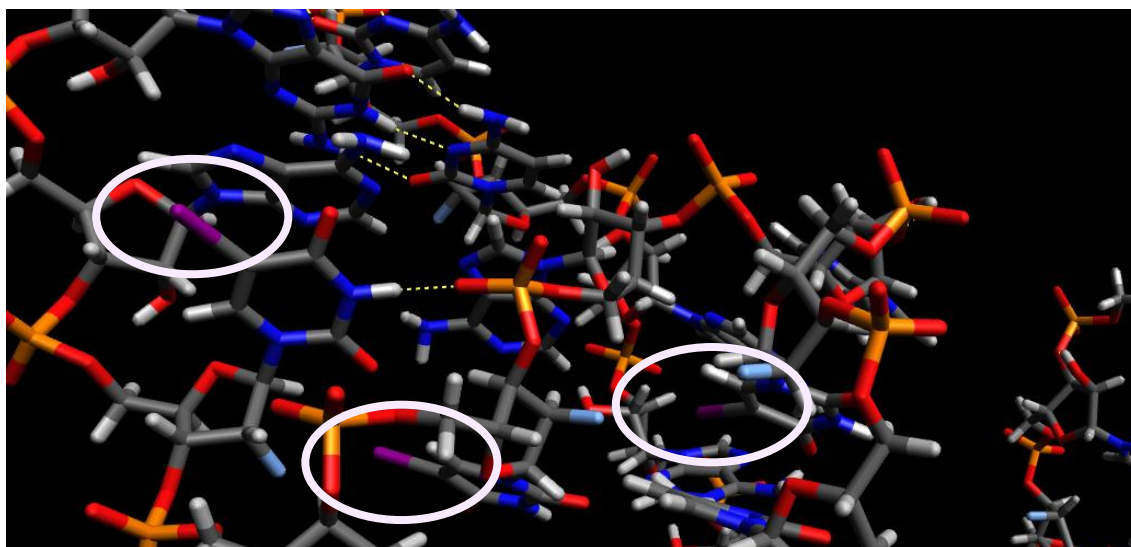
Figure 4.51. Iodo edit Zoom in of 3D Predicted secondary structure of MinE07U-I aptamer from Avogadro.
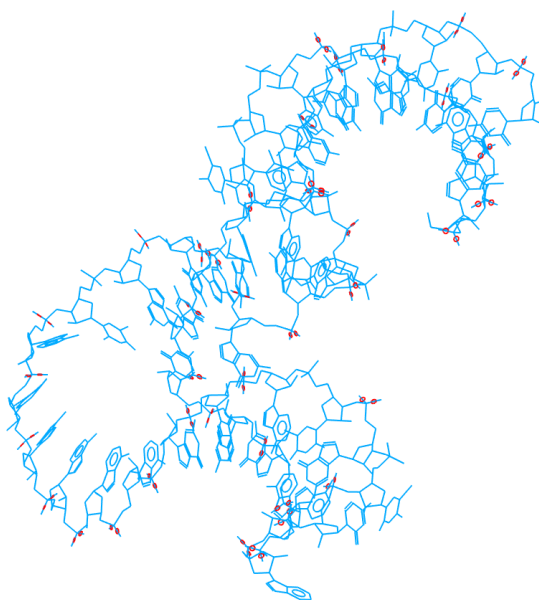


Figure 4.52. 3D Predicted secondary structure of MinE07-U-I aptamer aptamer in Avogadro with optimised

geometry using force field MMFF94.
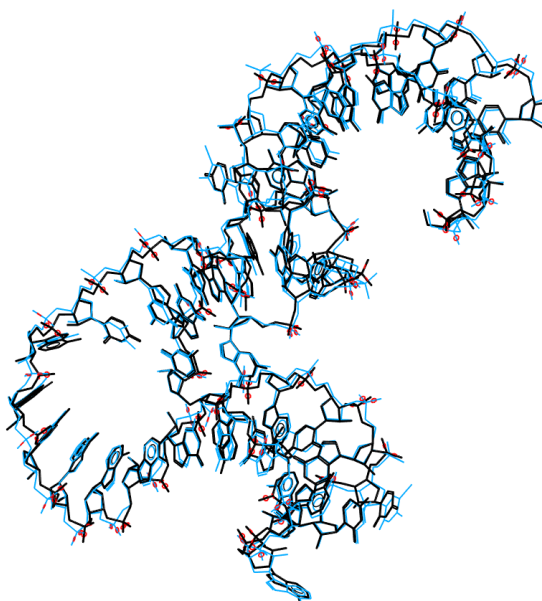
Figure 4.53. 3D Predicted secondary structure of MinE07-U-I aptamer (Blue) overlaid with MinE07 aptamer (Black). There are no black boxes on this figure as there are not any obvious differences.

Changing the hydrogen atom to an iodine does not change the conformation of the aptamer a huge amount as it does not require much more space, so aptamer **MinE07-U-I** visually looks similar to parent aptamer **MinE07** (4.53).

### 4.4.4.4. Comparison of the 3D Predicted Structures of MinE07, MinE07U-Ph, MinE07U-Vi, and MinE07U-I

The 3D predicted secondary structures of **MinE07, MinE07-U-Ph, MinE07-U-Vi** and **MinE07-U-I** were generated so that visual analysis could be undertaken to see what the modifications did to the folding and shape of the original parent **MinE07** (figure 4.54 and 4.55).
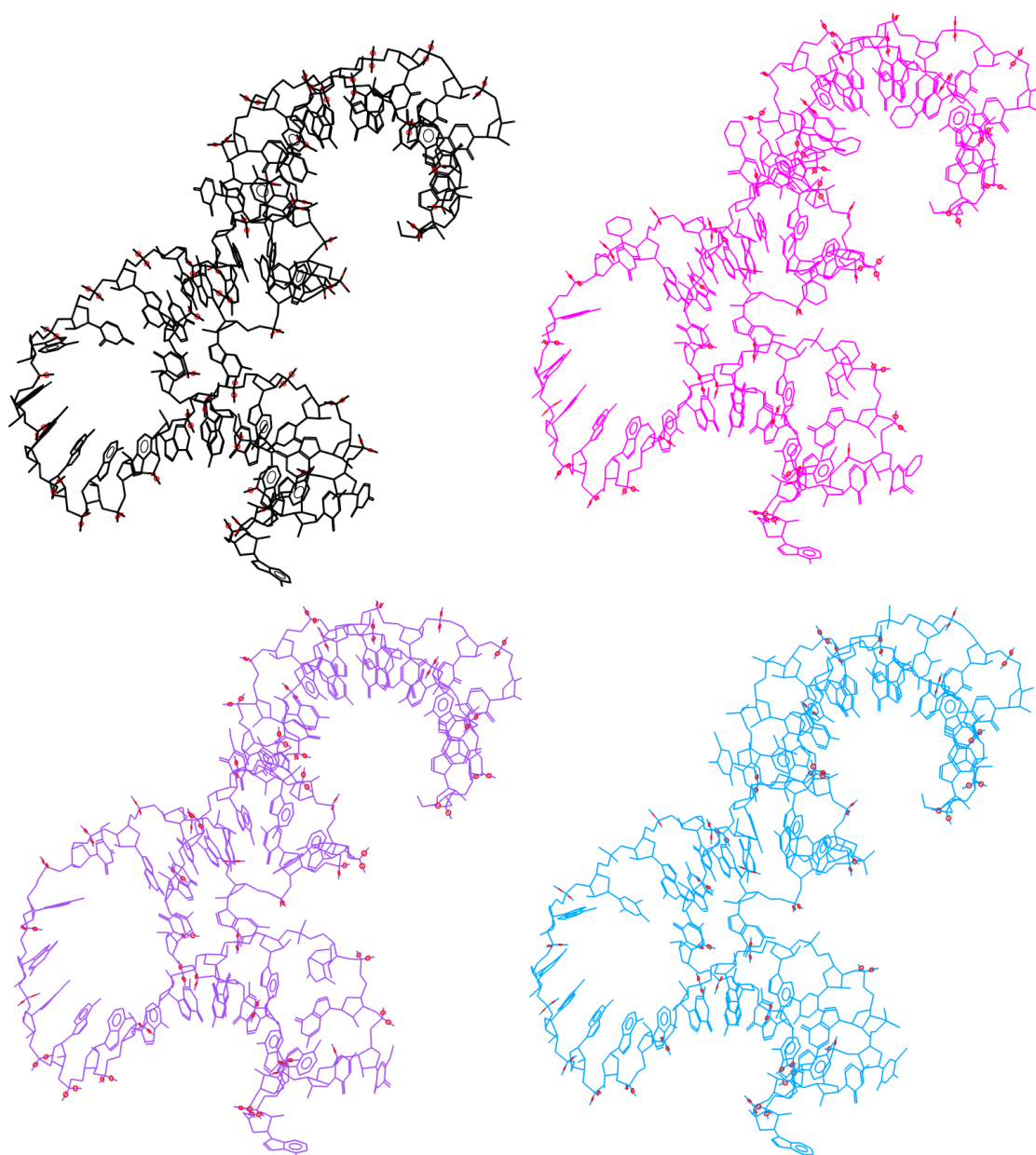
Figure 4.54. 3D Predicted secondary structure of MinE07 (Black), MinE07-U-Ph (Pink), MinE07-U-Vi (Purple) and

MinE07U-I aptamer (Blue).
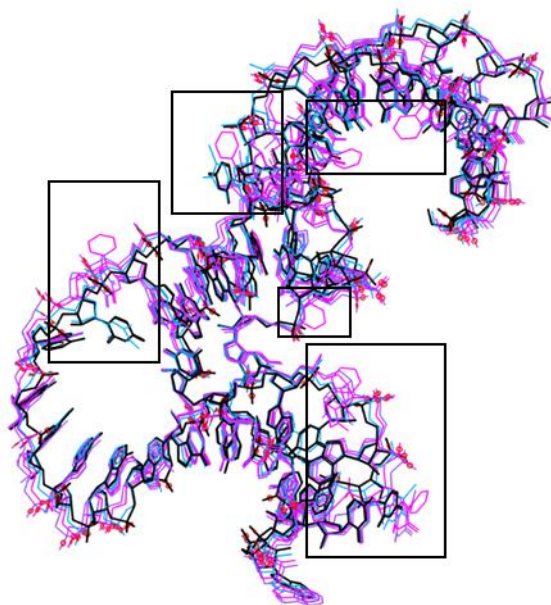
Figure 4.55. 3D Predicted secondary structure of MinE07 (Black) overlaid with MinE07-U-Ph (Pink), MinE07-U-Vi (Purple) and MinE07U-I aptamer (Blue). The black boxes highlight the key 5 areas that are changed the most over the 4 different structures.

The aptamer that visually shows the biggest different in conformation is **MinE07-U-Ph**. The size of the Ph causes some of the nucleotides to have to rotate in another direction to fit in the space the Watson-Crick bonding has created. Aptamer **MinE07-U-I** visually looks most similar to parent aptamer **MinE07**. Changing a hydrogen atom to an iodine does not change the conformation of the aptamer a huge amount as it does not require much more space. Changing the shape of **MinE07** could potentially positive or negative in terms of effecting the binding affinity to EGFR. This will be experimentally explored using protein binding affinity assays (chapter 6).

### 4.4.4.5. Purification of MinE07-U-Ph-Biotin, MinE07-U-Vi-Biotin and MinE07-U-I-Biotin

The overall percentage yields of the aptamers could have been better however, enough was synthesised to be purified and used in protein affinity assays. These samples were purified by PAGE, cutting out the correct length bands to get a single sequence sample. These purified samples were then desalted with Zetadex resin and then run again on an analytical

polyacrylamide gel to check that the PAGE purification (figure 4.56). This gel potentially shows

they have been purified successfully, but to be sure they will be run on a HABA assay to check

of the biotin group on the 5' end.
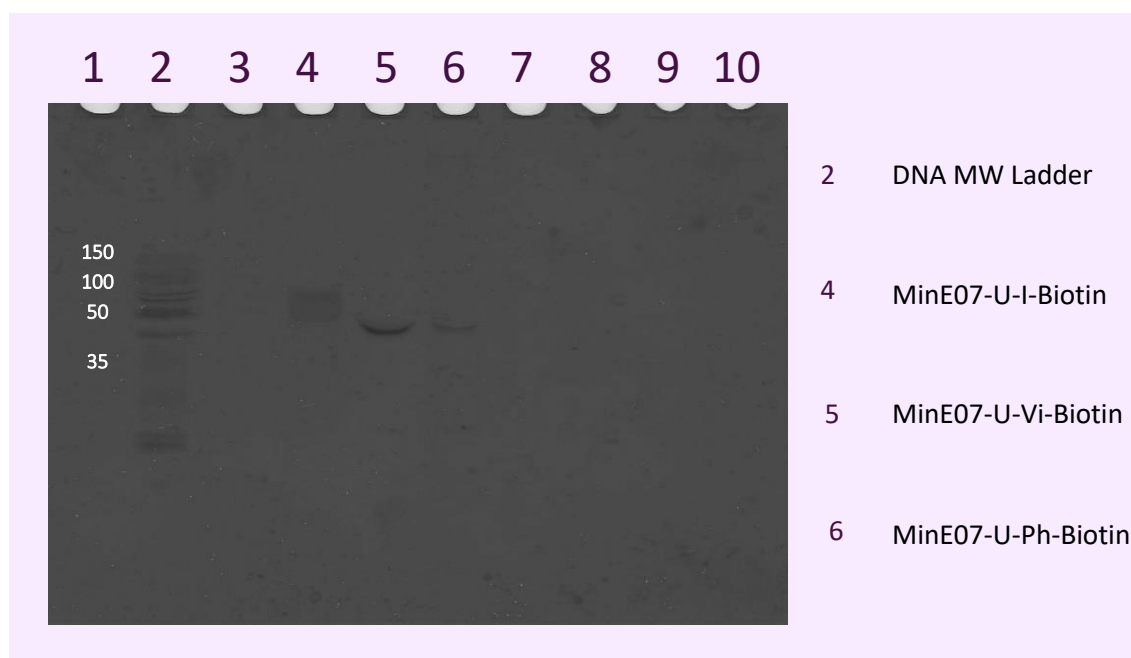


Figure 4.56. 15% Polyacrylamide Gel of MinE07-U-Ph-Biotin, MinE07-U-Vi-Biotin and MinE07-U-I-Biotin after PAGE

purification.

## 4.5. Conclusion

In conclusion, the aptamer synthesis was optimised for the successful synthesis of the MinE07

aptamer library, the control aptamers and the fully modified aptamers **MinE07-Biotin,**

**MinE07M3-Biotin, MinE07M6-Biotin, MinE07-U-Ph-Biotin, MinE07-U-Vi-Biotin** and

**MinE07-U-I-Biotin** Aptamers. The yields of these were high enough to carry on to the sorting

stage for the MinE07 aptamer library and the HABA/ protein binding assays for the fully

modified aptamers.

The computational aptamer modelling gives a visual of how these modifications effect the

shape and the folding of the **MinE07** aptamer highlighted by the black boxes on the figures.

Future work would include running some computational docking studies to work out the exact

position(s) of residues in the **MinE07** that dictate the high affinity of this aptamer for the EGFR.

# Chapter 5

Selection and Identification of High Affinity Modified MinE07 Aptamers using Flow Cytometry and Mass Spectrometry

# Chapter 5

## 5.1. Introduction

### 5.1.1. Aptamer Selection

#### 5.1.1.1. High throughput screening

High throughput screening (HTS) is the use of automated equipment to rapidly test up to millions of samples for biological activity at either the cellular, model organism, pathway, or molecular level.[464] A common HTS experimental process is when $10^3$ -$10^6$ small molecule compounds of known structure are screened in parallel. Other substances have also been screened this way such as oligonucleotides, antibodies, natural product extracts and chemical mixtures. Paul Ehrlich and Sahachirō Hata were among the first to screen agents to identify those that killed parasites grown in culture. The results of their experiments informed library screening for the next century.[465,466] After Merrifield's result on solid-phase peptide synthesis it brought HTS into mainstream industrial practice.[467] HTS aims to screen $10^6$ or more samples per day and because of this the assay designs are relatively simple and automation compatible. They aim to have robotic-assisted sample handling and automated data processing. HTS is commonly used in the biotechnology and pharmaceutical industries to identify compounds (hits) with biological or pharmacological activity. These hits are then used as a starting point for medicinal chemical optimisation during pharmacological probe or drug discovery and development.[468,469] Before HTS, screening was considerably more time consuming and expensive. In the past decade there has been a large increase in the number of new biological targets that are screenable as a result of the development of assay design and improvements in genomics and bioinformatics.[470] HTS assays are performed in plates with 96-, 384 or 1536 wells, and will

traditionally test each compound in a compound library at a single concentration, commonly 10 μM. Quantitative high throughput screening (qHTS) is a method of testing compounds at multiple concentrations at once using an HTS platform. This can then generate a concentration response curve for each compound tested straight as part of the screen. Recently, qHTS has become popular in toxicology because it can fully characterise the biological effects of chemicals for efficiency and decrease the rates of false positives and negatives.[471]
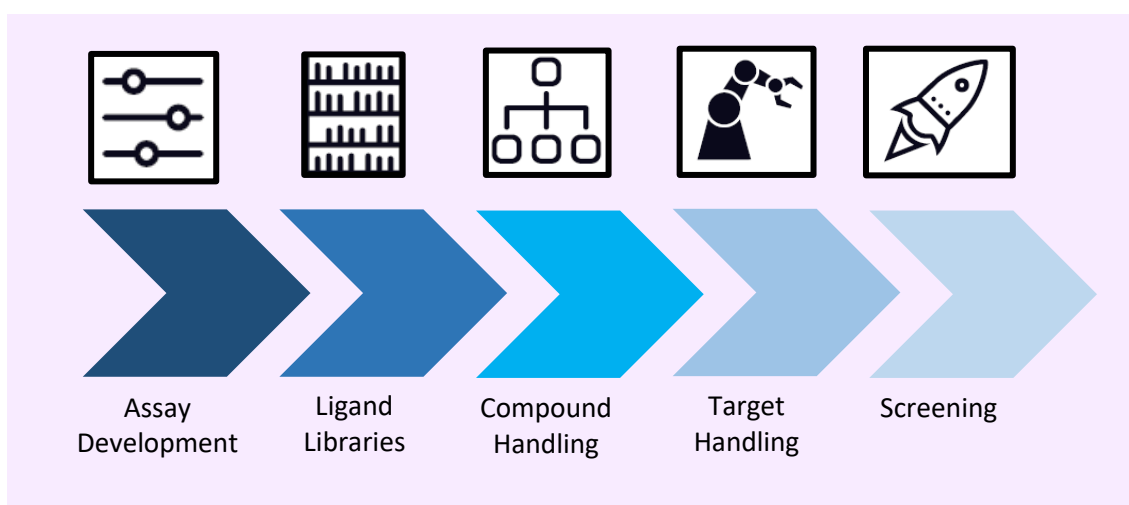


Figure 5.1. High-through put screening workflow.

### 5.1.1.2. HTS by Flow Cytometry

In the field of drug discovery flow cytometry is becoming an increasing popular tool for HTS assays for identifying new potential drug candidates.[472] Flow cytometers can be equipped with HTS attachments for automated sample collection into multi-well plates. The benefits for using flow cytometry for HTS assays are: (i) samples libraries can be screened rapidly, using millions of cells or compounds in a day. (ii) Many different cell lines or protein types can be screened singly or simultaneously, and in-depth analysis can be performed on both suspension and adherent cells (iii) HTS using flow cytometry, can also be used as a diagnostic tool.

### 5.1.1.3. Protein A

Protein A-FITC is used in protein binding affinity assays as it has a fluorescence tag and was chosen here because of its strong affinity to the Fc tag attached to the EGFR and for its low affinity to all other components in the assays. Protein A is a 42 kDa surface protein (figure 5.2). It was originally found in the cell wall of the bacteria *Staphylococcus aureus*. It is encoded by the *spa* gene and its regulation is controlled by DNA topology, cellular osmolarity and a two-component system known as ArlS-ArlR.[473] It is is composed of five homologous Ig-binding domains that fold into a three-helix bundle. Each domain is able to bind proteins from many mammalian species. It binds to the heavy chain within the Fc region of most immunoglobulins and so is a useful tool for antibody detection or purification.[474]
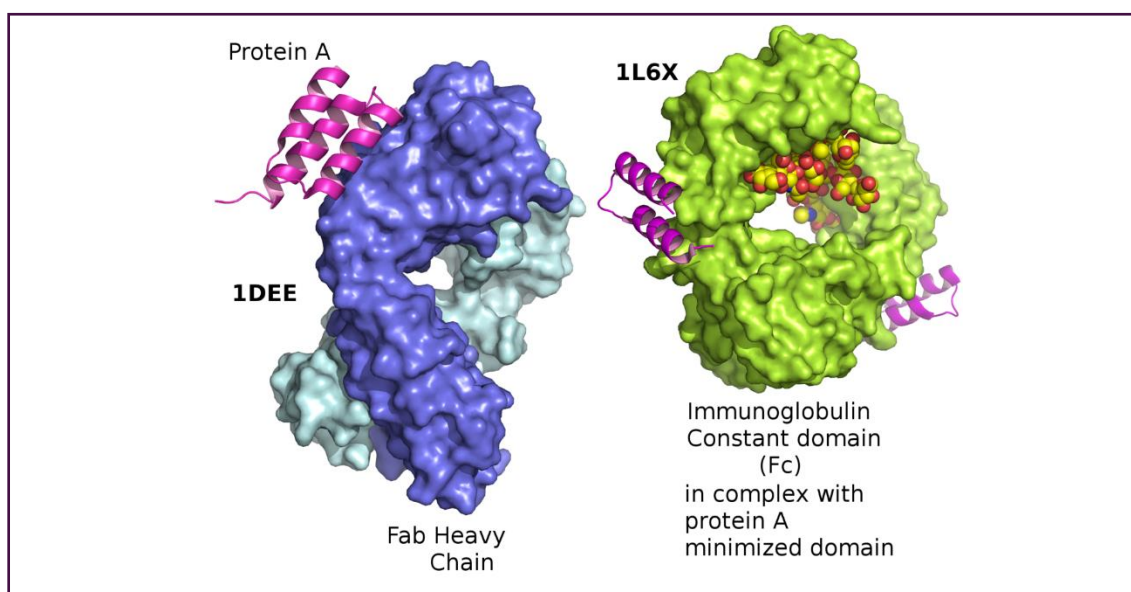


Figure 5.2. Protein A bound to antibodies

### 5.1.1.4. Streptavidin

Streptavidin is a 52.8 kDa protein (figure 5.3). It is purified from the bacterium *Streptomyces avidinii*.[475] It has an extremely high affinity for biotin, with a dissociation constant of ~$10^{-15}$ M. Streptavidin is used extensively in bio-nanotechnology and molecular biology for techniques

such as western blotting,[476] ELISA assays[477] and conjugation strategies.[478] This is because It is a high affinity biotin binding agent, it is resistant to extreme pHs, organic solvents, temperature, proteolytic enzymes and detergents.[479] The binding of biotin to streptavidin is one of the strongest non-covalent interactions known in nature. Streptavidin will be used as a tool in the protein binding assays, because of its very high affinity for biotin, which is bound to the end of the MinE07 modified aptamers that will be tested.
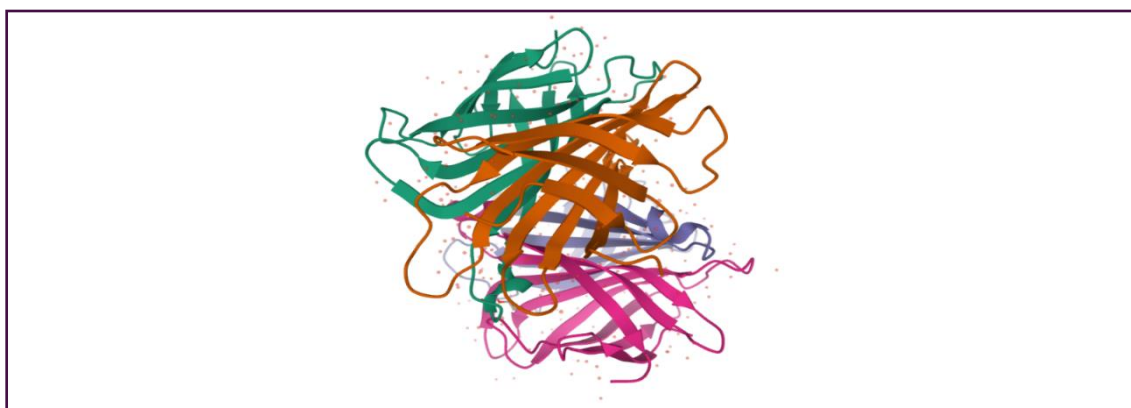


Figure 5.3. Crystal Structure of Streptavidin Protein.[480]

### 5.1.1.5. Biotin

Biotin is a small organic molecule with the molecular weight of 244.31 g/mol (figure 5.4). Biotin also known as vitamin H, also known as vitamin B$_7$ is a water-soluble B vitamin. B vitamins play an important role in cell metabolism. It is an essential component of enzymes involved in metabolising fat and carbohydrates, promoting cell growth and influencing amino acids that are participating in the synthesis of proteins.[481] Biotin binds very tightly to tetrameric protein avidin, it is one of the strongest known protein-ligand interactions.[482] The relatively simple chemical conjugations of biotin to a wide range of targets make it an attractive compound to work with.
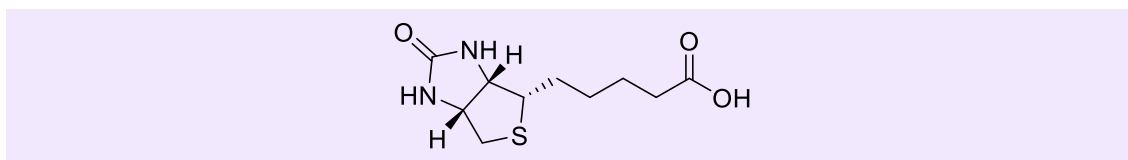
Figure 5.4. The chemical structure of Biotin

### 5.1.1.6. Biotinylation

Biotinylation is the process of binding biotin covalently to a protein, nucleic acid or other molecule. Biotinylation is a rapid reaction and is not likely to disrupt the naturally occurring function of biotin. Biotin binds to avidin and streptavidin with high specificity and affinity (figure 5.5). The binding of streptavidin/avidin to biotin is resistant to proteolysis, heat and changes in pH. This has allowed the use of biotinylated molecules in a large range of environments and conditions.[483]
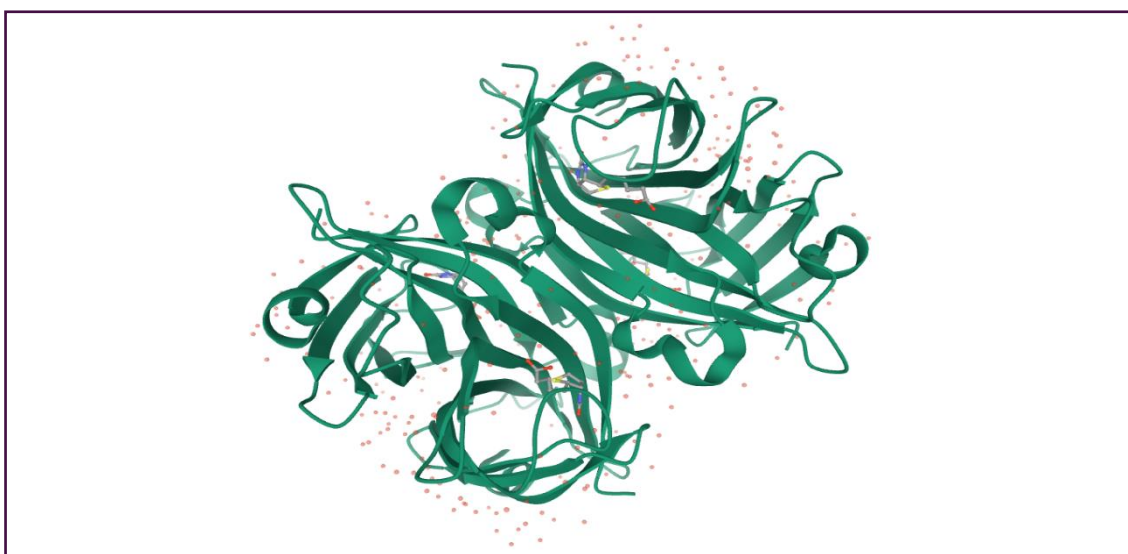


Figure 5.5. Streptavidin bound with Biotin[484]

### 5.1.1.7. Fc-Tag
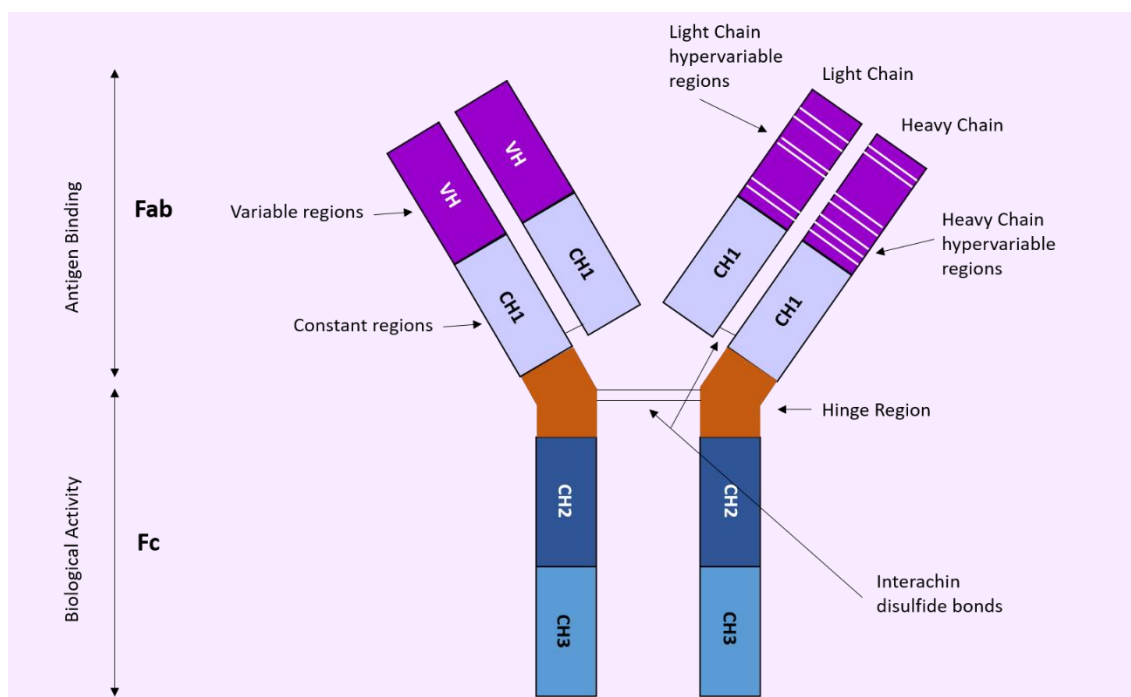


Figure 5.6. Antibody Structure showing the Fc tag.

Fc-tag is derived from the immunoglobulin Fc domain (figure 5.6). The fragment crystallisable region (Fc) is the tail region of an antibody that interacts with cell surface receptors called Fc receptors and some proteins of the complement systems. This property allows antibodies to activate the immune system[485].
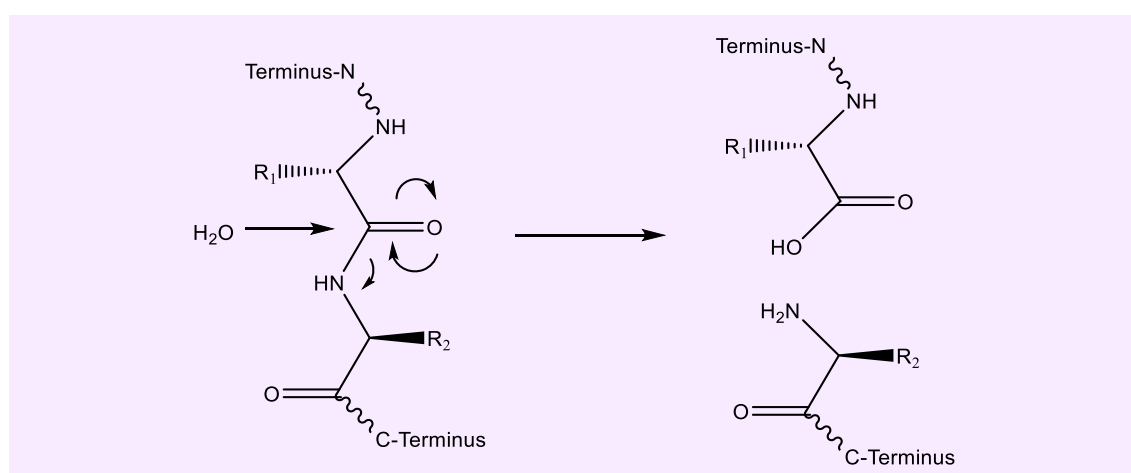


Figure 5.7. Process of proteolytically cleaving antibodies to form Fab and Fc fragments.

Antibodies also known as 'immunoglobulins' are glycoproteins (class of protein which have carbohydrate groups attached to the polypeptide chain) that bind antigens with high specificity and affinity through interactions such as electrostatic, hydrophobic, hydrogen bonds and Van der Waals forces. In humans there are 5 chemically and physically distinct classes of antibodies IgG, IgA, IgM, IgD and IgE. Antibodies have a basic unit of four polypeptide chains – two identical pairs of light (L) chains and heavy (H) chains which are bound together by covalent disulphide bridges as well as by noncovalent interactions. These molecules can be proteolytically cleaved to yield two Fab fragments (the antigen-binding part of the molecules) and the Fc fragment (the part of the molecule responsible for effector functions). Both the H and L chains are divided into V and C regions. The V region contains the antigen-binding site and the C region determines the fate of the antigen. Fc will be used as a tool in the protein affinity assays, as the recombinant EGFR protein has an Fc tag at the Fc has a high affinity for Protein A which is bound to the 96-well plates being used in the binding assay.

## 5.1.2. Tandem Mass Spectrometry

Tandem mass spectrometry also known as MS/MS is a technique in instrumental analysis where two or more mass spectrometers are coupled together using an additional step to increase their ability to analyse chemical samples.[486] Tandem mass spectrometry is most commonly used for the analysis of proteins and peptides. The molecules of a given sample are ionized and the ions are separated by their mass-to-charge (m/z) ratio by the first spectrometer. Ions of a particular m/z ratio are selected and then made to split into small fragment ions by either collision-induced dissociation, ion-molecule reaction or photodissociation. These fragments are then introduced into the second mass spectrometer which separates the fragments by their m/z ratio and detects them. The fragmentation step allows for the identification of sequences such as amino acids.

### 5.1.2.1. Locating Chemical Modifications through LC-MS Based Analysis

Knowledge of the structural information is essential for understanding the functional details of modified RNA.[487] Liquid chromatography coupled with mass spectrometry (LC-MS) analysis provides direct and sensitive detection of each and every chemical modification in RNA (figure 5.8). This platform relies upon the physico-chemical properties of modified nucleosides such as altered chromatographic and mass spectrometric behaviour as well as changes in mass/charge values. This method was pioneered by the McCloskey group.[488] MS methods also allow for the identification of modified RNA by collision induced dissociation-based tandem mass spectrometry.[489]

Complete hydrolysis of RNA to nucleosides would lose the information about the specific positions at which modifications occur. It is therefore better to have partial hydrolysis through chemical or enzymatic treatments which can then provide sequences that can help better identify where the modifications occur. Acid[490] or alkali[491] treatment allows for random cleavage of the phosphodiester backbone, which then generates oligonucleotides that are of appropriate length for MS sequencing through collision induced dissociation (CID). These oligonucleotides can show sequence overlap which helps determine the overall sequence. Chemical hydrolysis has limitations such as potential degradation of the RNA or modifications because of the extreme pH conditions, poor control on the extent of cleavage, and that the RNA needs a high level of purity. Modification positions can be preserved by treatment with nucleobase specific ribonucleases. These oligonucleotide digestion products vary in the length of fragments they produce, depending on the position of the recognised nucleobases in the sequence. This selective RNA cleavage method has the advantage of restricting the compositional value, decreasing the number of potential base compositions for a given mass measurement.[489]
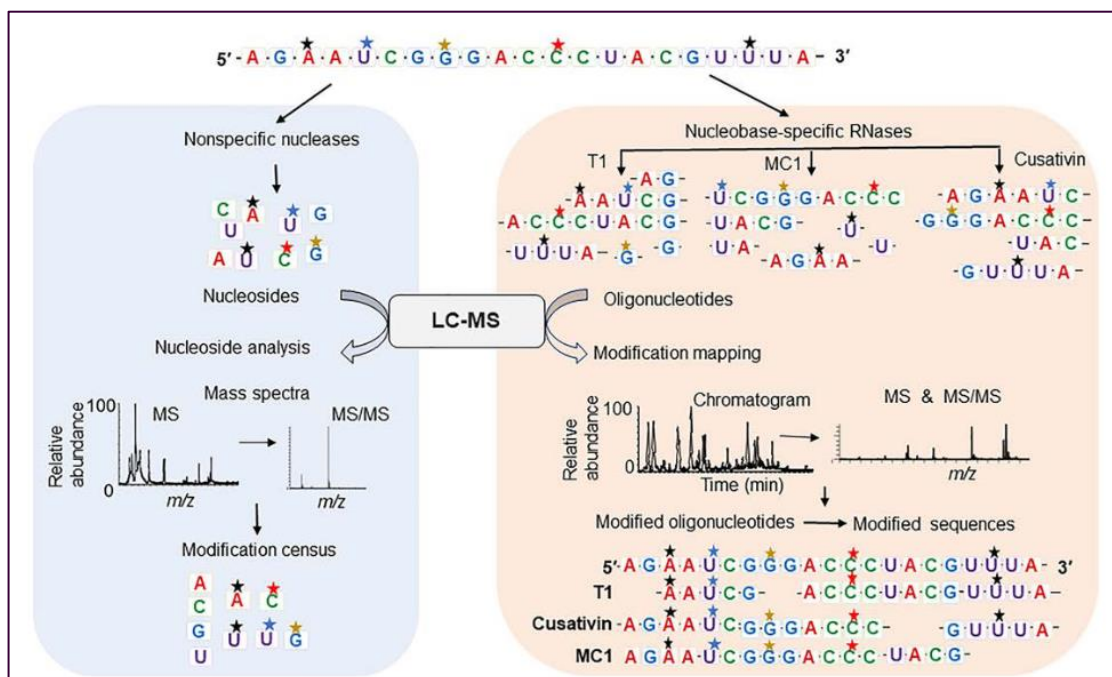
Figure 5.8. Characterisation of chemical modifications in RNA through LC-MS/MS analytical platform. The left-hand side depicts the workflow for identification of nucleoside modifications. The right-hand side depicts the workflow used for locating the positions of chemical modifications in the RNA sequence which involves digestion with nucleobase-specific ribonucleases and oligonucleotide sequencing. Chemical modifications are shown by a star.[487]

The RNA sequence and its modifications can be mapped to the specific sites by monitoring the mass shift or increase in mass of oligonucleotide molecular ion and its sequence information in the fragment ions during tandem mass spectrometry methods (figure 5.9). The interpreted oligonucleotide sequences are matched against the target/parent RNA sequence to assign the locations and identity of each modification.
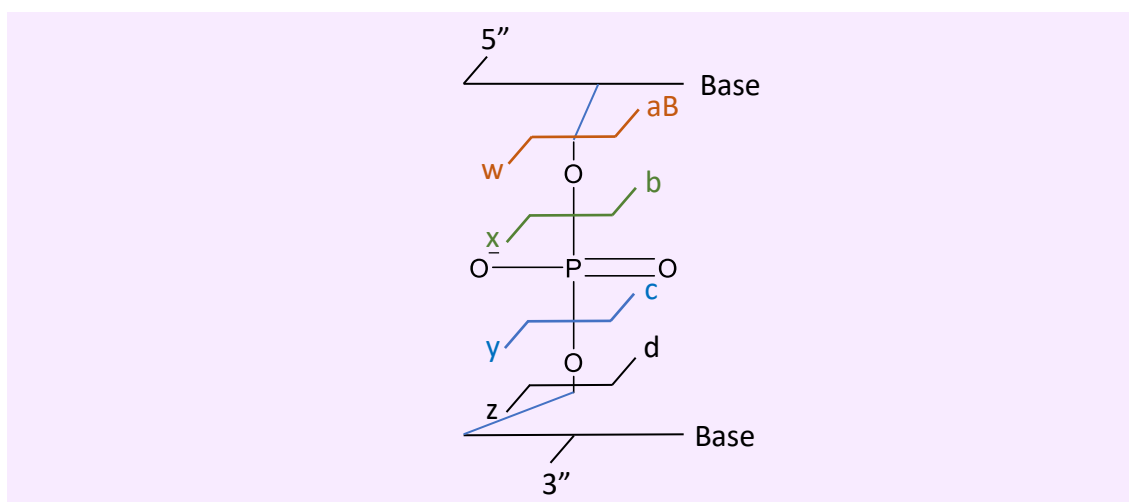
Figure 5.9. Collision-induced dissociation of phosphodiester backbone in an oligonucleotide. Cleavage of P - O bond followed by C – O bond are predominant in RNA which leads to generation of complementary fragment ions c and y and a-B and w.[487]

### 5.1.2.2. Analysis of LC-MS/MS Data and RoboOligo Software

RNA, tRNA and other biological or synthetic RNA polymers can contain nucleotide modifications. Traditional Sanger sequencing methods[492] cannot establish the chemical nature and sequence of these modified-nucleotide containing oligomers. MS has become the conventional approach for determining the nucleotide composition, modification status and sequence. Modified RNAs are analysed by MS using CID tandem mass spectrometry (CID MS/MS), which produces a complex set of data of oligomeric fragments. These fragments must be interpreted to identify the location of the modification within the RNA sequence.

RoboOligo was developed in 2015, as an interactive software program for the robust data analysis generated by CID MS/MS of RNA oligomers. Software has been developed to make the data analysis easier, as samples run on the MS/MS produce a large amount of complicated data. The three main functions of RoboOligo are: (i) automated *de novo* sequencing via the local search paradigm (ii) manual sequencing with real-time spectrum labelling and cumulative intensity scoring (iii) a hybrid approach, called 'variable sequencing', which combines the user

intuition of manual sequencing with the high-throughput sampling of automated *de novo* sequencing.[493] CID of oligoribonucleotides typically produces c-, y-, w- and a-B-type product ions (figure 5.10), although, any bond within the phosphodiester backbone is susceptible to dissociation.[494] The goal during MS/MS data analysis of modified oligonucleotides is to use the detected product ions to reconstruct the original sequence.[494] Successive product ions coming from the same phosphodiester backbone fragmentation (c1-, c2-, etc.) help determine the nucleotide ordering within the oligoribonucleotide. Modified nucleosides are placed within the oligoribonucleotide sequence by their unique nucleoside residue mass which is normally greater than the residue mass of the four original ribonucleosides.[489] The MS/MS data produced by MALDI or LC-ESI-MS is difficult to analyse due to the complexity of the resulting spectra and that there are minimal available software tools to help with the analysis.[495]



Figure 5.10. Typical RNA oligomer fragmentation products generated by CID. The most abundant fragment ions are usually c- and y- fragments. W- and a-b- ions are often also observed.

The first attempt as computational analysis of mass spectrometric data focused on the determination of the nucleotide composition of an ion based on the mass of the oligomer and the masses of the four original ribonucleosides.[496] The Simple Oligonucleotide Sequencer (SOS) was released in 2002 by Rozenski and McCloskey. SOS was capable of assisting the manual interpretation of oligonucleotide MS/MS data from up to 20 base length oligomers.[497] SOS worked by displaying the mass spectral peaks a-B- or w- fragment ions as the user matching them with the corresponding nucleoside. This program was effective but limited by the manual

256

analysis of data preventing its use on complex data sets and also a minimal number of modifications could be evaluated during analysis. More recently in 2013, Nyakas *et al.*[498] developed OMA and OPA programs. These allowed the analysis of MS and MS/MS data with a customisable database of nucleotides. The software only compares the predicted fragmentation pattern with an inputted sequence, as the limitation is that the sequence has to be known in advance. Two database search strategies for RNA MS data: RRM[499] and Ariadne.[500] They have been developed to analyse the data in a similar way to the polypeptide analysis suite MASCOT other similar protein-focused software.[501] RRM focuses on mass spectral data only and compares it with a database of RNA sequences, so is not effective with modified nucleosides. Ariadne scores the comparison of the data to an inputted database of theoretical ribonuclease digested and CID-fragmented RNA sequences. This software can currently only be used for sequences from *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Mus musculus* and *Homo sapiens*. Automated *de novo* sequencing of MS/MS data has been attempted using various different methods from samples of DNA[502], DNA adducts[503] and RNA containing 2 -*O*-methyl and phosphorothioate linkages.[504]

RoboOligo was developed to fill in some of the limitations of the other programs, creating an interactive program equipped with the ability for robust analysis of negative ion mode MS/MS data generated by CID. They show that an automated local search model maintains the robustness of the global search model and can successfully handle modified bases. There is additional data analysis flexibility, which is provided by manual and variable sequence capabilities that allow for examination of MS/MS data by the user.[493]

## 5.1.3. Chapter Aims

The aim of this chapter is to design and optimise the aptamer selection and identification method. It will be achieved using flow cytometry along with tandem mass spectrometry. This method is novel as aptamers have not before been selected and identified together in this way.

# 5.2. Materials and Instrumentation

## 5.2.1. Materials

TentaGel® M NH$_2$ Monosized Amino TentaGel Microspheres were purchased from Rapp Polymere. Sheath Fluid, 8 peak calibration bead and Accudrop beads were purchased from BD Biosciences. Recombinant Human EGFR Fc Chimera Protein, CF was purchased from R&D Systems. Native *Staphylococcus aureus* Protein A (FITC) (ab7455) was purchased from Abcam. White PS – Protein A coated –level 200 µL 96 well plates were purchased from Biomat. Costar assay 96 well plates (with lid, white, flat bottom, tissue culture treated) were purchased from Thermo Fisher Scientific. Mass spectrometry columns nanoE MZ Sym C18 Trap Column 5 µm and nanoE MZ HSS T3 Column 1.8 µm 75 µm x 150 mm were purchased from Waters. Methanol hypergrade for LC-MS LiChrosolv®, Water for chromatography (LC-MS Grade) LiChrosolv® and MicroSpin™ G-50 columns were purchased from Sigma Aldrich.

### 5.2.1.1. Buffers

Wash Buffer: PBS + 0.05% Tween 20, to 500 mL of PBS $^{(+/+)}$, add 250 µL of Tween-20.

Protein Binding Buffer: 48.5 mL of PBS $^{(+/+)}$, 1 mL 50 mg/mL BSA and 0.5 mL 10 mg/mL ssDNA were added.

## 5.2.2. Instrumentation

### 5.2.2.1. Fluorescent Activated Cell Sorter

Specification can be found in Chapter 2 section 2.2.2.2 Fluorescent Activated Cell Sorter

### 5.2.2.2. Mass Spectrometer

Specification can be found in Chapter 3 section 3.2.2.3 Mass Spectrometer

### 5.2.2.3. Plater Reader

All 96 well plates were scanned with a Victor X4 plate reader from Perkin Elmer using either fluorescence or chemiluminescence settings available.

### 5.2.2.4. UV-Visible Absorption

Specification can be found in Chapter 2 section 2.2.2.5 UV-Visible Absorption Spectra

# 5.3. Experimental

**5.3.1. Checking the stability of the Rhodamine B tagged TentaGel® M NH$_2$ Monosized Amino TentaGel Microspheres**

**5.3.1.1. Running the Rhodamine B tagged TentaGel® M NH$_2$ Monosized Amino TentaGel Microspheres through the FACS**

The FACSJazz™ lasers were calibrated before the samples were run with the 8 peak BD beads (standard calibration). Samples TGRhodB100.2 TGRhodB50.2 TGRhodB25.2 TGRhodB10.2 TGRhodB1.2 were run through the FACS dispersed in 5mL sheath fluid, being observed on laser 488 nm with filter 585/29 nm. All data was collected and analysed using the BD FACS Sortware sorter software program.

### 5.3.1.2. Running the 6-carboxyfluorescein tagged TentaGel® M NH₂ Monosized Amino TentaGel Microspheres through the FACS

Standard calibration was performed (5.3.1.1). Samples TGCFluro100.1 TGCFluro50.1 TGCFluro25.1 were run through the FACS dispersed in 5 mL sheath fluid, being observed on laser 488 nm with filter 513/17 nm. All data was collected and analysed using the BD FACS Sortware sorter software program.

### 5.3.2. Protein binding assay to check for non-specific binding with EGFR-Fc and Protein A-FITC with TentaGel® M NH₂ Monosized Amino TentaGel Microspheres

A plain 96 well plate was washed with wash buffer (150 µL x 3). Sheath fluid (100 µL), TentaGel® M NH₂ Monosized Amino TentaGel Microspheres in sheath fluid (100 µL) and Protein A-FITC (100 µL of 1 µg/mL) were added to the appropriate wells as background signals (100 µL per well). Four samples were made up. Sample 1 was TentaGel microspheres (0.0034 g) incubated with binding buffer (300 µL) for 30 minutes on a shaker at 300 rpm. The microspheres were spun down and the binding buffer was removed and the microspheres were washed with wash buffer. EGFR-Fc (300 µL of 1 µg/mL) was added and incubated together with the microspheres for 60 minutes on a shaker at 300 rpm. The microspheres were spun down EGFR-Fc was removed, and the microspheres were washed with wash buffer. Protein A-FITC (300 µL of 1 µg/mL) was added to the microspheres and incubated together for 60 minutes on a shaker at 300 rpm. This was spun down and the supernatant was put in wells and the beads were re-suspended in water and added to separate wells. Sample 2 was TentaGel microspheres (0.0034 g) incubated with binding buffer (300 µL) for 30 minutes on a shaker at 300 rpm. The binding buffer was removed, and the microspheres were washed with wash buffer. Wash buffer (300 µL) was added instead of EGFR-Fc and incubated with the microspheres for 60 minutes on a shaker at 300 rpm. The microspheres were spun down and the buffer was removed, and the microspheres were washed with wash buffer. Protein A-FITC (300 µL of 1 µg/mL) was added

and incubated together with the microspheres for 60 minutes on a shaker at 300 rpm. This was spun down and the supernatant was put in wells and the beads were re-suspended in water and added to separate wells. Sample 3 was MinE07 tagged TentaGel microspheres (MinE07ARP) (0.0034 g) incubated with binding buffer (300 μL) for 30 minutes on a shaker at 300 rpm. The microspheres were spun down and the binding buffer was removed, and the microspheres were washed with wash buffer. EGFR-Fc (300 μL of 1 μg/mL) was added to the microspheres and incubated together for 60 minutes on a shaker at 300 rpm. The microspheres were spun down EGFR-Fc was removed, and the microspheres were washed with wash buffer. Protein A-FITC (300 μL of 1 μg/mL) was added to the microspheres and incubated together for 60 minutes on a shaker at 300 rpm. This was spun down and the supernatant was put in wells and the beads were re-suspended in water and added to separate wells. Sample 4 was MinE07 tagged TentaGel microspheres (MinE07ARP) (0.0034 g) incubated with binding buffer (300 μL) for 30 minutes on a shaker at 300 rpm. The binding buffer was removed, and the microspheres were washed with wash buffer. Wash buffer (300 μL) was added to the microspheres instead of EGFR-Fc and incubated together for 60 minutes on a shaker at 300 rpm. The buffer was removed, and the microspheres were washed with wash buffer. Protein A-FITC (300 μL of 1 μg/mL) was added to the microspheres and incubated together for 60 minutes on a shaker at 300 rpm. This was spun down and the supernatant was put in wells and the beads were re-suspended in water and added to separate wells. The 96 well plate that these samples were added to was then read on the Victor X4 plate reader after 10 minutes (Fluorescein (485 nm/535 nm, 1.0 s).

### 5.3.3. Protein binding assay to check for non-specific binding with EGFR-Fc and Protein A-FITC with TentaGel® M NH₂ Monosized Amino TentaGel Microspheres analysed on the FACS.

The microsphere samples from experiment 5.3.2 were diluted in 4 mL of sheath fluid and placed in FACS falcon tubes for analysis. Standard calibration was performed as described in 5.3.1.1.

The 488 nm laser was used with filter 513/17 nm. All data was collected and analysed using the BD FACS Sortware sorter software program.

### 5.3.4. Two-way sort of 100% 6-carboxyfluorescein tagged microspheres from plain microspheres.

6-carboxyfluorescein tagged microspheres (1 mL) and plain microspheres (1 mL) were added together in a FACS flacon tube and sonicated. Standard calibration was preformed (5.3.1.1). The 6-carboxyfluorescein tagged microspheres were sorted into the right tube (red gate) and the plain microspheres were sorted into the left tube (blue gate). Waste was collected into another falcon tube. The collected samples were put back through the FACS for pre-sorting analysis. The 488 nm laser was used with filter 513/17 nm. All data was collected and analysed using the BD FACS Sortware sorter software program.

### 5.3.5. Two-way sort of Accudrop Beads to confirm photo-bleaching theory

Accudrop beads (2 mL) were put into a FACS sized falcon tube. Standard calibration was performed (5.3.1.1). The drop-delay value was adjusted while viewing BD FACS Accudrop beads in the centre and side sort streams that are illuminated by a red diode laser. Sort monitoring was undertaken with live video feed of breakoff point, waste collection, and side streams. The accudrop beads were sorted into two gates P7 and P8. The beads that were sorted into P7 were then put back through the FACS and sorted again in gates P7 and P8. This was repeated again. The 488 nm laser was used with filter 513/17 nm. All data was collected and analysed using the BD FACS Sortware sorter software program.

### 5.3.6. Two-way sort of 100% 6-carboxyfluorescein tagged microspheres from plain microspheres to confirm photo-bleaching theory

Standard calibration was performed (5.3.1.1). 6-carboxyfluorescein tagged microspheres (2 mL) were put into a FACS sized falcon tube. They were sorted into two gates P7 and P8. The

microspheres that were sorted into P7 were then put back through the FACS and sorted again in gates P7 and P8. This was repeated again. The 488 nm laser was used with filter 513/17 nm

### 5.3.7. 96 well plate sort of 100% 6-carboxyfluorescein tagged microspheres

Standard calibration was performed (5.3.1.1). The FACS was calibrated using the 96 well plate set up program along with the accudrop beads. 6-carboxyfluorescein tagged microspheres (2 mL) were put into a FACS sized falcon tube. 1 gate was created for the population of microspheres. The microspheres were sorted into wells B1-9, C1-9, D1-9 in a ½ log digression starting at 50,000 microspheres down to 5 microspheres. The 488 nm laser was used with filter 513/17 nm. All data was collected and analysed using the BD FACS Sortware sorter software program. The 96 well plate samples were then read on the Victor X4 plate reader (Fluorescein (485 nm/535 nm, 1.0 s).

### 5.3.8. Sorting MinE07Lib to extract the top binding MinE07 modified aptamers

All of the MinE07Lib (1 µM, 0.0029 g) was added to an Eppendorf tube. Binding buffer (100 µL) was added and the tube was incubated at room temperature for 15 minutes. This was centrifuged at 300rpm for 1 minute; the binding buffer was removed and the microspheres were then washed with binding buffer (2 x 100 µL). EGFR-Fc (0.03357 mg/mL, 100 µL) was then added to the tube and this was incubated at room temperature for 60 minutes. This was spun down at 300 rpm for 1 minute, the EGFR-Fc was removed, and the microspheres were then washed with wash buffer (2 x 100 µL). Protein A-FITC (0.03357 mg/mL, 100 µL) was then added to the tube and this was incubated at room temperature for 60 minutes. This was spun down at 300 rpm for 1 minute, the Protein A-FITC was removed, and the microspheres were then washed with wash buffer (2 x 100 µL). Sheath fluid (4 mL) was added, and the microspheres sonicated. The sample was then added to a FACS sized falcon tube. Standard calibration was performed (5.3.1.1). The MinE07Lib-EGFR-Fc-Protein-A-FITC microspheres sample was two-way

263

sorted three times, changing the gates each time to match the photobleaching and to select

less each time. The FACS was calibrated using the 96 well plate set up program along with the

accudrop beads. The top microspheres selected from the first three rounds of sorting the

MinE07Lib-EGFR-Fc-Protein-A-FITC microsphere were then sorted across two 96 well plates

with 100 µL of RO water. The 488 nm laser was used with filter 513/17 nm. All data was collected

and analysed using the BD FACS Sortware sorter software program.

### 5.3.9. Photocleaving the top hit MinE07Lib aptamers

Both 96 well plates containing the top MinE07Lib hits in was placed under a UV lamp for 3 hours

to photocleave the aptamers from the TentaGel microspheres. The contents of each well (100

µL) was then transferred into individual PCR tubes and stored at 4˚C.

### 5.3.10. LC-MS/MS of MinE07-Biotin

Mobile phase A is 8 mM tetraethyl ammonium bromide (TEAB) in LC-MS grade water adjusted

to the pH 7.5-7.8. Mobile phase B is 8 mM TEAB in a 1:1 ratio of LC-MS grade methanol and

water. The MinE07-Biotin samples are dissolved in 10 µL of LC-MS water to the concentrations

of 1 µM, 0.5 µM, 0.2 µM, 0.1 µM, 0.05 µM, 0.02 µM and 0.01 µM. A was initially held at 92%

and B at 8%. B was increased to 65% at 30 mins and up to 95% at 31 mins. At 33 mins B was

dropped back to 8% and then held until 60 mins. The column was maintained at 50˚C.[505, 487]

### 5.3.11. LC-MS/MS of Hits from MinE07Lib Selection

The samples were purified with MicroSpin™ G-50 columns. Mobile phase A was 8 mM tetraethyl

ammonium bromide (TEAB) in LC-MS grade water adjusted to the pH 7.5-7.8. Mobile phase B

was 8 mM TEAB in a 1:1 ratio of LC-MS grade methanol and water. The MinE07Lib samples were

in nuclease free water at an unknown concentration. Initially A was held at 92% and B at 8%. At

30 mins B was increased to 65% and up to 95% at 31 mins. At 33 mins B was dropped back to

8% and then held until 60 mins. The column is maintained at 50˚C.[505, 487] 15 MinE07Lib hit were randomly chosen to be analysed.

### 5.3.12. Analysis of top Hit Aptamers LC-MS/MS data

The data generated by LC-MS/MS of the top hit aptamers MinE07-139, MinE07-41 and MinE07-56 was analysed using the program RoboOligo.[493]

# 5.4. Results and Discussion

## 5.4.1. Stability of the fluorescently tagged TentaGel® M NH₂ Monosized Amino TentaGel Microspheres

### 5.4.1.1. Checking the stability of the Rhodamine B tagged TentaGel® M NH₂ Monosized Amino TentaGel Microspheres

To assess how stable the fluorescently tagged TentaGel microspheres are, samples TGRhodB100, TGRhodB73, TGRhodB36, TGRhodB17 and TGRhodB1.7 were run through the FACS, being observed on laser with the excitation of 561 nm and the emission of 585/29 nm (same as chapter 2.4.3.2). The FACS data shows that the Rhodamine B tagged TentaGel microspheres are therefore stable in sheath fluid at room temperature for several years (experiments being run in 2019 and 2021) (figure 5.11). Having a solid support which is stable over a long period of time is important if several rounds of aptamer selection need to be performed.
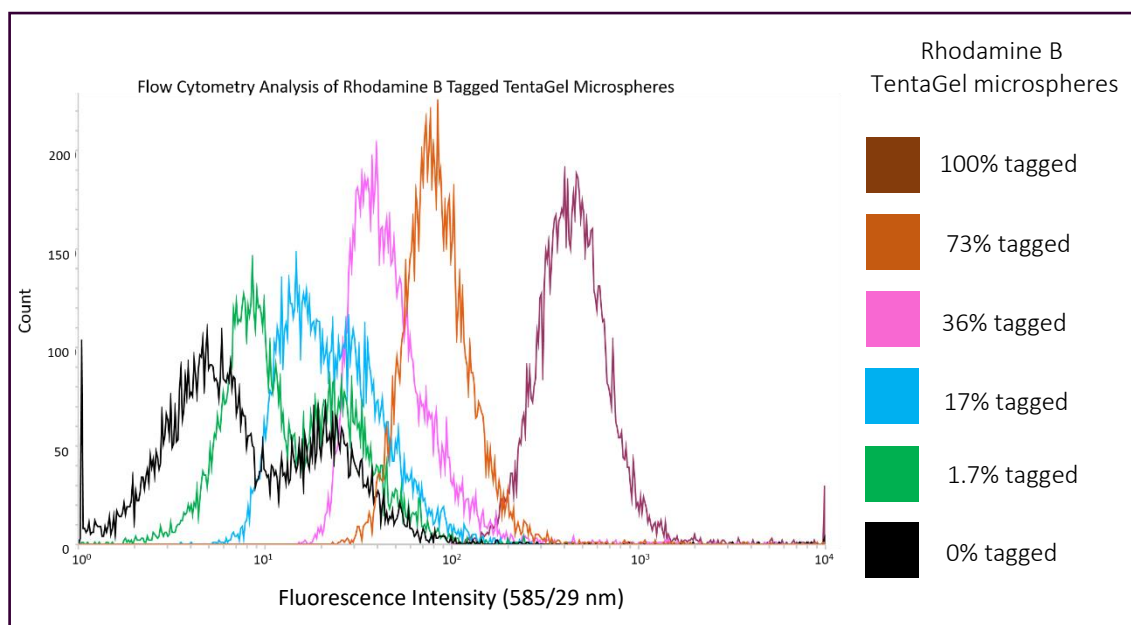
Figure 5.11. Flow Cytometry Analysis of Rhodamine B TentaGel Microspheres. Doublets caused by two microspheres going through the laser at the same time.

### 5.4.1.1. Checking the stability of the 6-carboxyfluorescein tagged TentaGel® M NH₂ Monosized Amino TentaGel Microspheres

The same process was performed with the 6-carboxyfluorescein tagged TentaGel microspheres. Samples TGCFluor100, TGCFluor73 and TGCFluor36 were ran through the FACS dispersed in 5 mL sheath fluid, being observed on laser 488 nm with filter 513/17 nm (chapter 2.4.3.2). The FACS data shows that the 6-carboxyfluorescein tagged TentaGel microspheres are stable in sheath fluid at room temperature for several years (figure 5.12). There is not show such a large different in the 73% and 100% tagged microspheres, and this fluorescent data is not as successful.
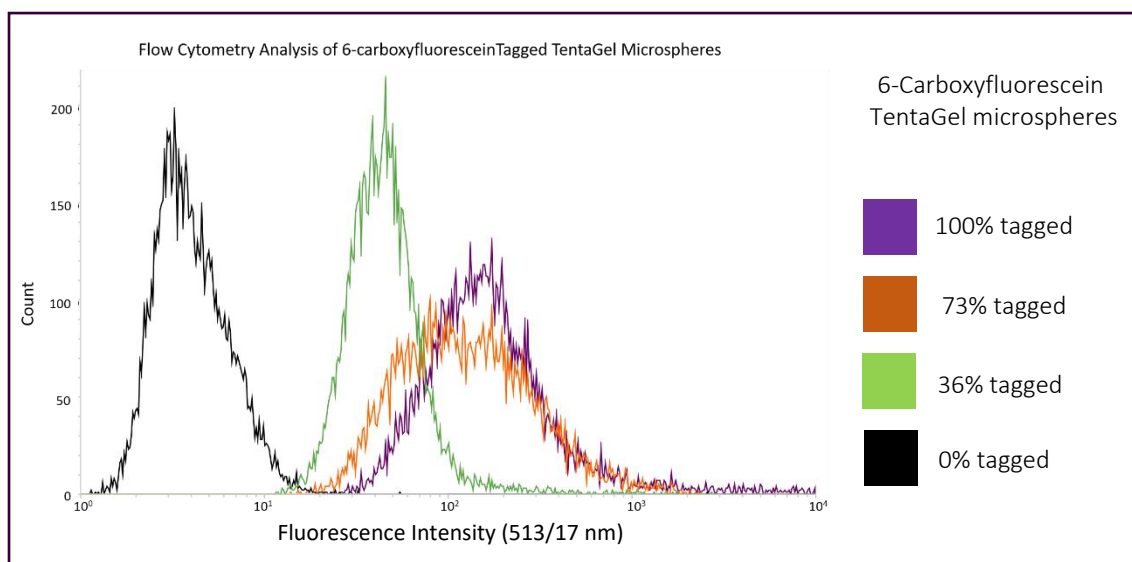
Figure 5.12. Flow Cytometry Analysis of 6-carboxyfluorescein TentaGel Microspheres.

## 5.4.2. Non-specific binding of assay reagents with TentaGel® M NH₂ Monosized Amino TentaGel Microspheres
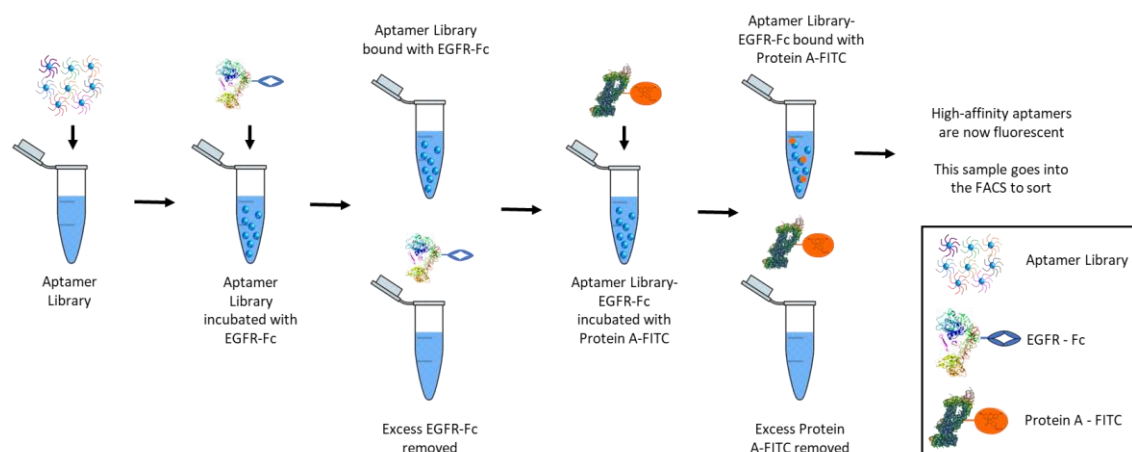


Figure 5.13. Schematic of the Protein Binding Chain to the Aptamer Library. The aptamer library first binds with EGFR-FC. The FC tag has high affinity for Protein A which has a fluorescent tag FITC. This allows the aptamers that bind to EGFR to become fluorescent.

A protein binding assay was performed to check for any non-specific binding between the assay reagents and the MinE07 tagged TentaGel microspheres. Previous protein binding assays have been run and confirmed that there is no non-specific binding between MinE07 and Protein A-FITC by Centauri Therapeutics (unpublished), however none had been done between the

267

TentaGel microspheres and EGFR-Fc or Protein A-FITC. Any non-specific binding would not allow for an accurate aptamer selection on the FACS. The results show that the binding of the EGFR-Fc was successful with the **MinE07ARP** microspheres (62771 RFU) and that a limited amount of signal was in the supernatant of this reaction (17206 RFU) (figure 5.14). The **MinE07ARP** microspheres bound with just Protein A-FITC show there is no non-specific binding between the MinE07ARP microspheres and Protein A-FITC. This is shown by the supernatant having a similar RFU number to Protein A-FITC (39760) at 38835, meaning they stayed in solution and did not bind with the MinE07ARP microspheres. The related microspheres also show a low fluorescence at 11626 RFU which is similar to the plain microspheres in solution (13916 RFU). The plain microspheres bound with Protein A-FITC show no non-specific binding as the supernatant has a RFU of 33851, similar to the Protein A-FITC and the microspheres have a low RFU of 11328 showing the Protein A-FITC did not binding to the plain microspheres. The plain microspheres bound with EGFR-Fc and Protein A-FITC show a minimal amount of non-specific binding as the supernatant has a RFU of 23536 which is higher than expected. The microspheres have a lower RFU of 20318 showing the minimal amounts of EGFR-Fc did bind to the plain microspheres. To overcome this potential issue more washing steps will be needed, the plain microspheres stick together more than the microspheres that have had linkers bound to them (as seen by the doublet peaks in the FACS data), so this is probably the cause of this.
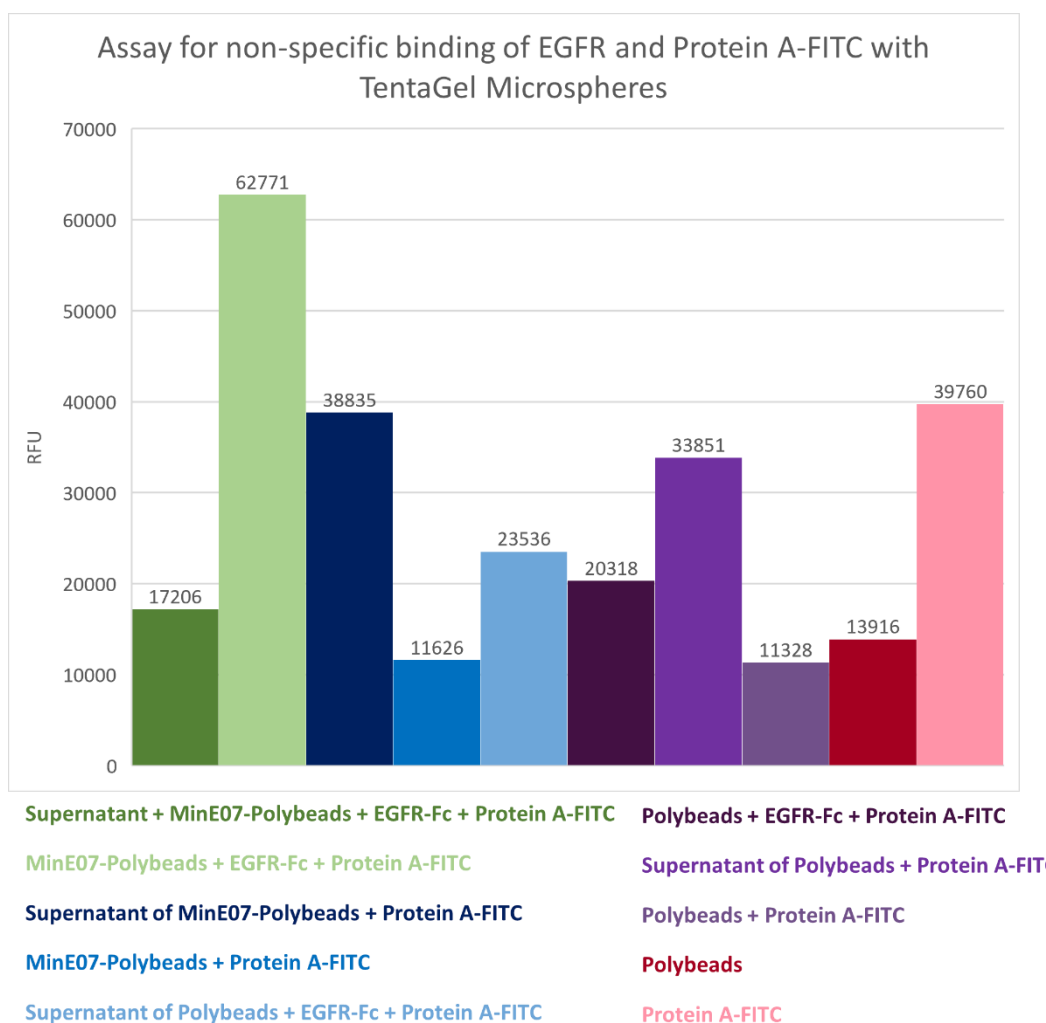
Figure 5.14. Data from the Assay for non-specific binding of EGFR and Protein A-FITC with TentaGel Microspheres.

These results were verified by running the samples through the FACS to check their fluorescence intensity using the 488 nm laser and the 513/17 nm filter. Figure 5.15 shows MinE07ARP-EGFR-Fc-Protein A-FITC Microspheres versus the plain microspheres. There is a large peak at $10^4$ which is the EGFR-Fc tightly bound to the **MinE07ARP** microspheres. There are some peaks lower down at $10^1$-$10^2$ which may be due to the synthesis of the **MinE07ARP** being the first successful synthesis of **MinE07** on these TentaGel beads and so not optimised. There is the potential for $1.14e^{11}$ aptamer strands to be synthesised on the surface of each microsphere. However, the full potential of the microspheres may not have been achieved during synthesis.
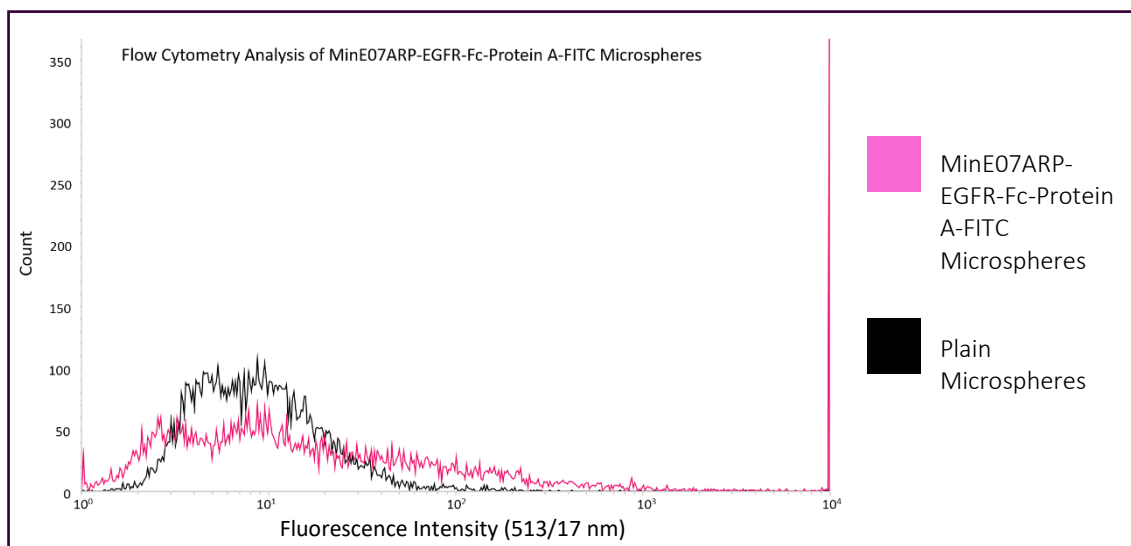
Figure 5.15. Flow Cytometry Analysis of MinE07ARP-EGFR-Fc-Protein A-FITC Microspheres.

All the microspheres from the non-specific protein binding assay were put through the FACS and compared with plain microspheres and 100% 6-carboxyfluorescein tagged microspheres. None of the microspheres show anything peaks above $10^2$ except for MinE07ARP-EGFR-Fc-Protein A-FITC and 100% 6-carboxyfluorescein tagged microspheres, which is the desired result (figure 5.16.). From the results of this assay and flow cytometry data the sorting selection will be gated at $10^4$.
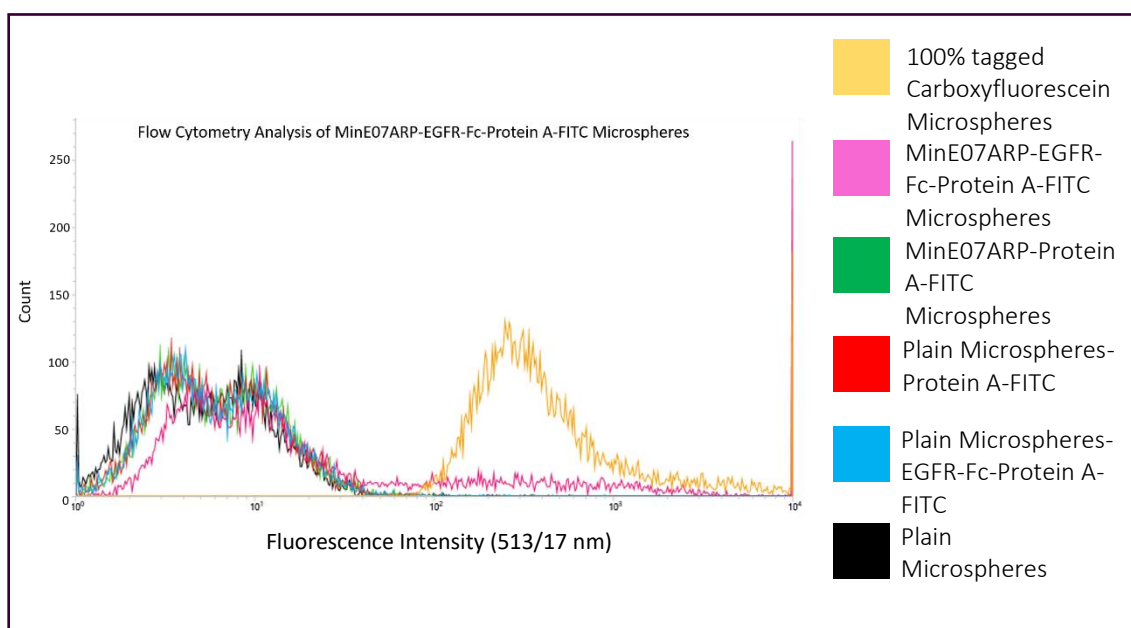


Figure 5.16. Flow Cytometry Analysis of MinE07ARP-EGFR-Fc-Protein A-FITC Microspheres.

## 5.4.3. Preliminary Sorting Experiments on the FACSJazz

5.4.3.1. Two-way sort of 100% 6-carboxyfluorescein tagged microspheres from plain microspheres

Preliminary sorting experiments were undertaken on the FACS to determine the best strategy for the sorting of the **MinE07** aptamer library, **MinE07Lib**. The first experiment performed was to sort 100% 6-carboxyfluorescein tagged microspheres from plain microspheres doing a two-way sort monitoring the fluorescence intensity using the 488 nm laser and the 513/17 nm filter. The dot plots show a clear difference in the two samples. The pre-sort data was used to determine the gates made to sort the sample (figure 5.17).
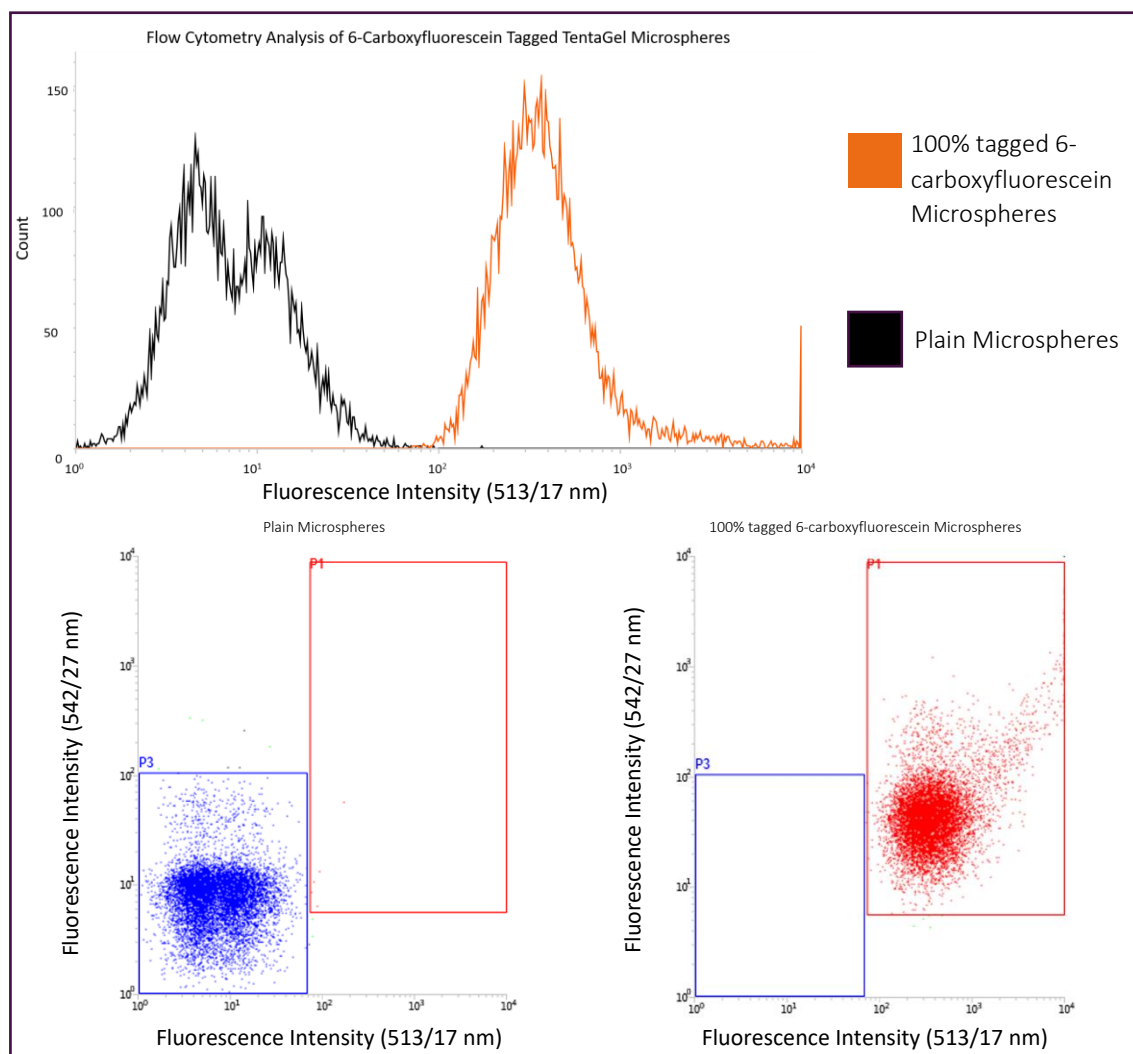


Figure 5.17. Flow Cytometry Analysis of 100% tagged 6-carbxyfluorescein microspheres.

The 6-carboxyfluorescein tagged microspheres were sorted into the right tube (red gate) and the plain microspheres were sorted into the left tube (blue gate). Waste was collected into another falcon tube. The data shows two distinct populations of microspheres (figure 5.18).
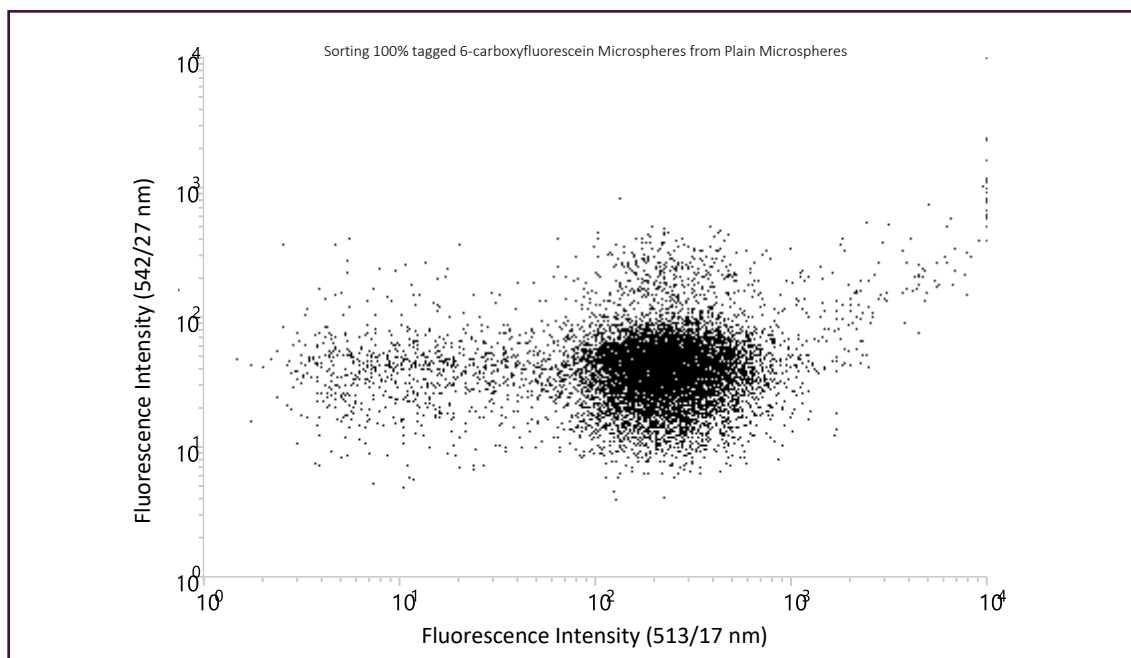


Figure 5.18. Flow Cytometry Analysis of 100% tagged 6-carbxyfluorescein microspheres.

The two-way sort samples collected were then put back through the FACS to check if the sort was successful. The data produced from re-analysing the sorted microspheres shows that some photobleaching was occurring to the fluorescently tagged microspheres. Figure 5.19 shows that the plain microspheres were in the same place as they were pre-sort however the 6-carboxyfluorescein microspheres were showing a lower fluorescence (red peak). The black peak and blue peak are similar whereas the orange peak and red peak are different whereas they should be the same.
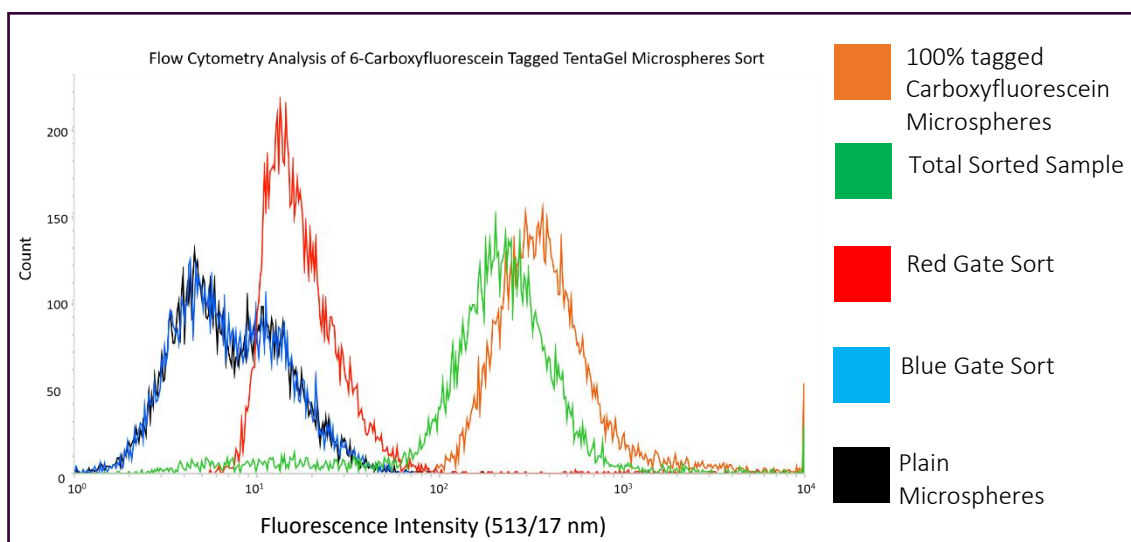
Figure 5.19. Flow Cytometry Analysis of 100% tagged 6-carbxyfluorescein microspheres after sort.

As the results from the two-way pre-sorting were not as expected another experiment was designed to confirm the theory of the fluorescent tags being photo-bleaching.

5.4.3.2. Two-way sort of 100% 6-carboxyfluorescein tagged microspheres from plain microspheres to confirm photo-bleaching theory

Accudrop beads that are used to calibrate the FACS before sorting were sorted two ways. All the beads that were collected in orange P7 gate were then put back through the FACS and re-sorted (figure 5.20). With each sort less beads were appearing in the orange P7 gate, showing that they were losing fluorescence intensity with each sort. In sort 1 purple P8 gate had 61% of the beads and orange P7 gate had 39% of the beads. In sort 2 purple P8 gate had 86% of the beads and orange P7 gate had 14% of the beads. In sort 3 purple P8 gate had 99% of the beads and orange P7 gate had 1% of the beads.
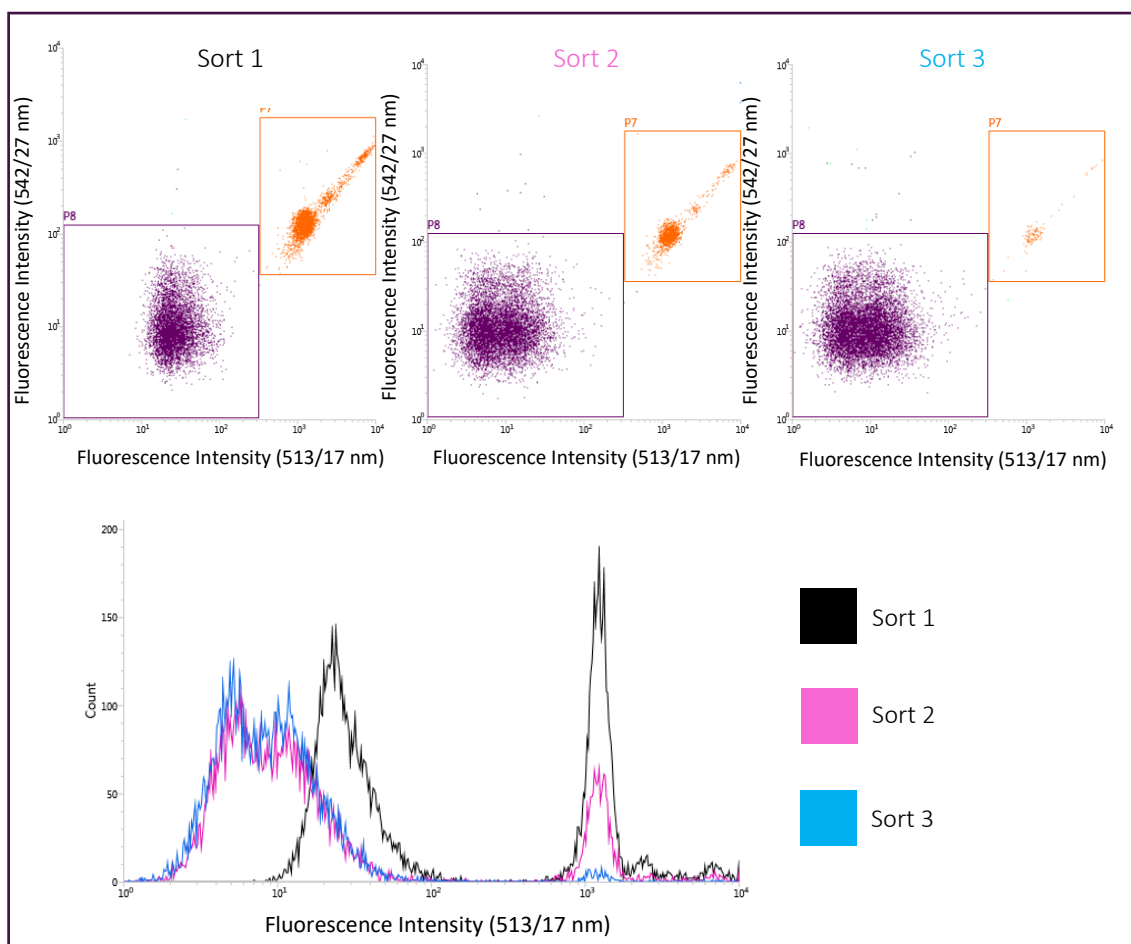
Figure 5.20. Flow Cytometry Analysis of Accudrop Beads Photo-bleaching.

The same experiment was then repeated with 100% tagged 6-carbxyfluorescein microspheres. The microspheres collected from orange P7 gate where then put through the FACS again. As with the Accudrop beads, with each sort the fluorescent intensity of the microspheres decreased each time, displaying a photo-bleaching effect (figure 5.21).
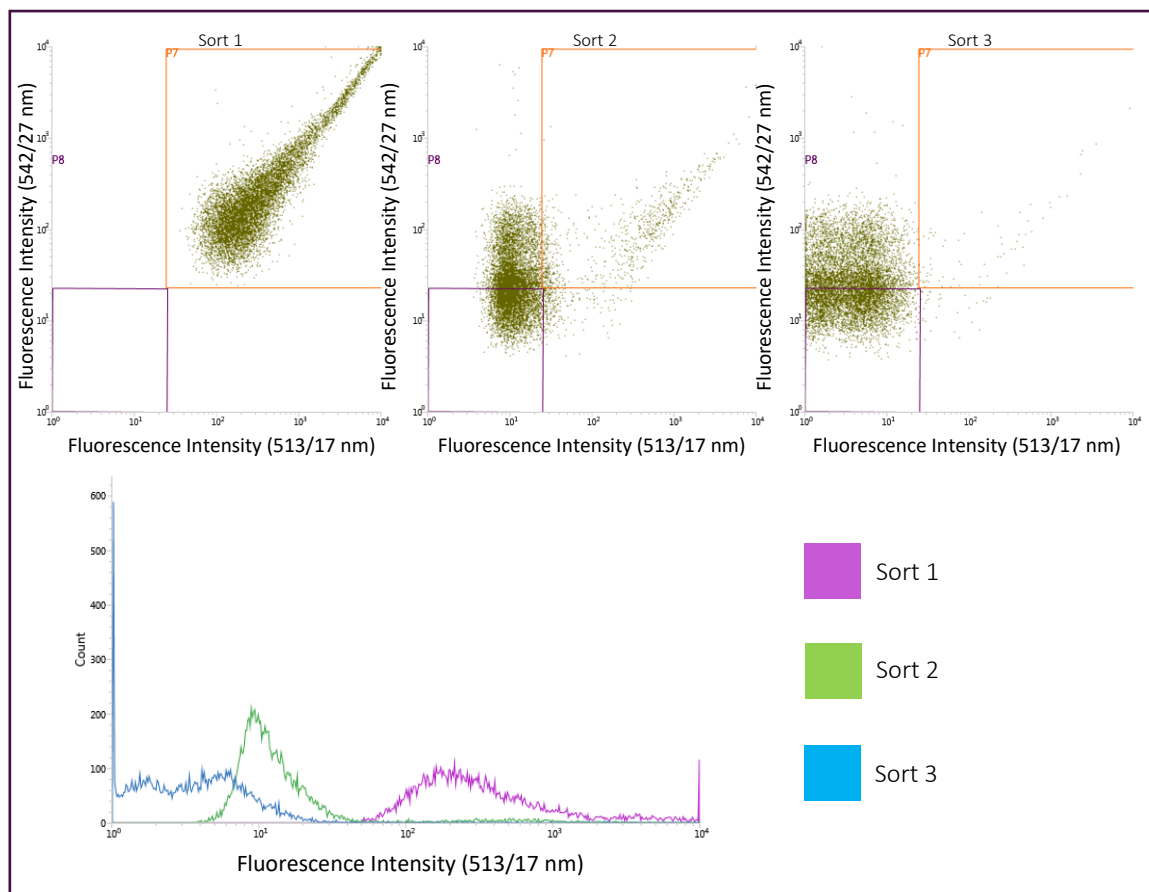
Figure 5.21. Flow Cytometry Analysis of 100% tagged 6-carbxyfluorescein microspheres Photo-bleaching.

### 5.4.3.3. Sorting MinE07ARP-EGFR-Fc-Protein-A-FITC microspheres from plain microspheres

To check that the two-way sort was sufficiently selecting the microspheres that were desired, sorting of high fluorescent MinE07ARP-EGFR-Fc-Protein A-FITC microspheres from a mix of MinE07ARP-EGFR-Fc-Protein-A-FITC microspheres and plain microspheres was attempted. From the analysis of MinE07ARP-EGFR-Fc-Protein-A-FITC microspheres 1.7% (385 microspheres) of the beads were in the orange P7 gate and 96.1% (21,390 microspheres) were in the purple P8 gate. If 1 mL of each sample was mixed together, the amount of microspheres sorted in orange P7 gate was predicted to be 0.85% of the total microspheres analysed and 98% in the purple P8 gate. The sort results show that 0.87% (199 microspheres) were sorted into the orange P7 gate and 98.1% (22,312 microspheres) were in the purple P8 gate (figure 5.22).
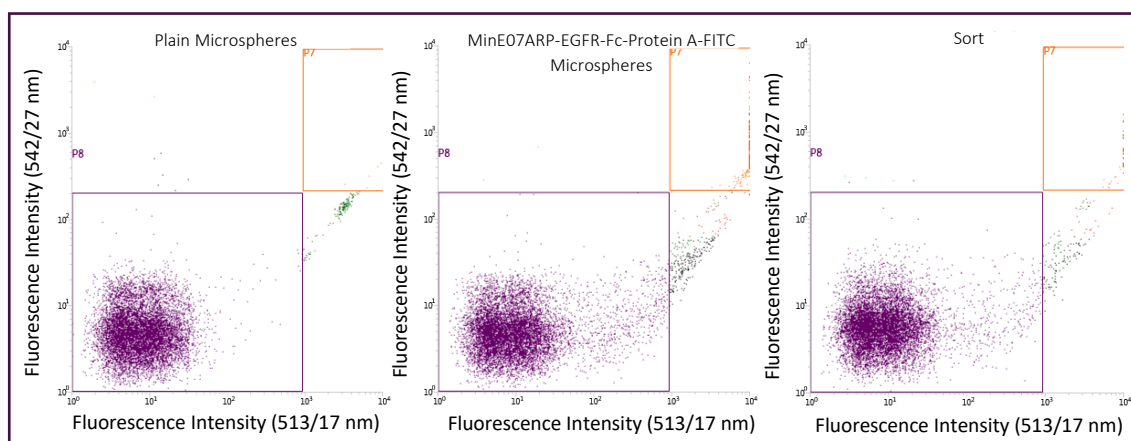
Figure 5.22. Flow Cytometry Analysis of MinE07ARP-EGFR-Fc-Protein A-FITC microspheres sorted from plain microspheres – 1st round.

This experiment was repeated with different gating, so as to count all the microspheres being put through the FACS (figure 5.23). From the analysis of MinE07ARP-EGFR-Fc-Protein-A-FITC microspheres 7.57% (5058 microspheres) of the beads were in the orange P7 gate and 92.5% (69,569 microspheres) were in the purple P8 gate. If 1 mL of each sample was mixed together the amount of microspheres sorted in orange P7 gate was predicted to be 3.75% of the total microspheres analysed and 96.25% in the purple P8 gate. The sort results show that 3.4% (2,278 microspheres) were sorted into the orange P7 gate and 96.6% (65,326 microspheres) were in the purple P8 gate. The results of these two sorting experiments prove that the desired microspheres can be sorted from a large population.
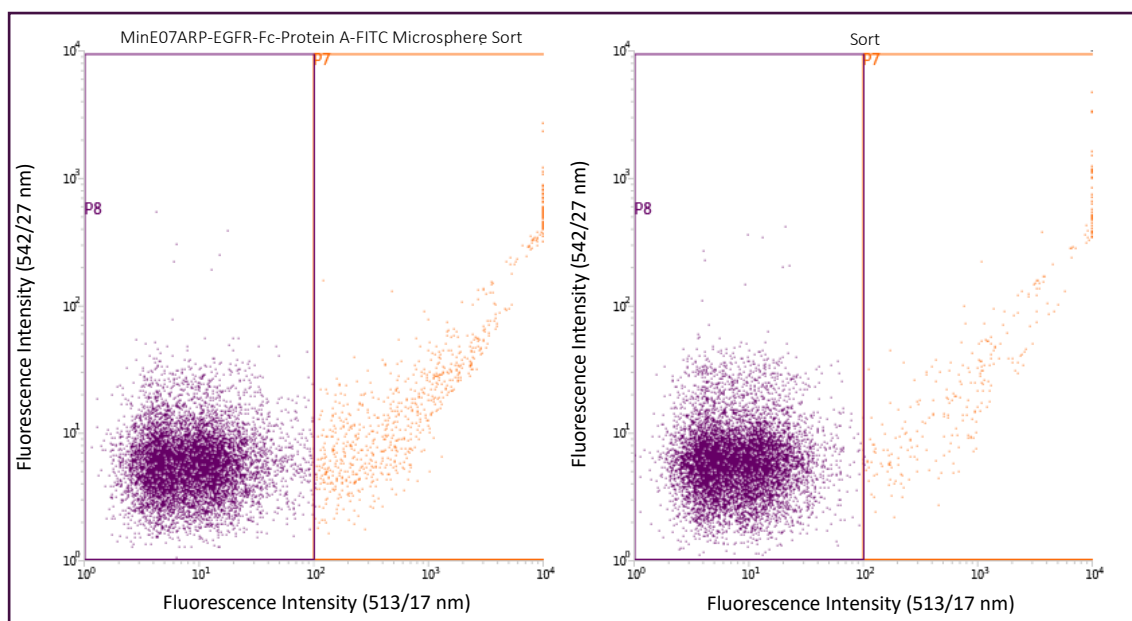
Figure 5.23. Flow Cytometry Analysis of MinE07ARP-EGFR-Fc-Protein A-FITC microspheres sorted from plain microspheres – 2nd round with new gating.

### 5.4.3.4. 96 well plate sort of 100% 6-carboxyfluorescein tagged microspheres

The next preliminary FACS sorting experiments that were undertaken were sorted into a 96 well plate, as it would be beneficial to have the top sorted microspheres all separate for the mass spectrum analysis. This experiment was done to check the FACS accurately separates out the microspheres in to each well. The 96 well plate was set up to sort in a ½ log dilution style, across a row in triplicate. The number of microspheres per well can be found in figure 5.24 along with the FACS sort data.

**Sort Report**

Sort Target Details

| Name | Population | Event Limit | Event Count | Sort Count | Sort Rate | Time (sec) |
|------|-----------|-------------|-------------|------------|-----------|------------|
| D1 | P7 | 50,000 | 50,191 | 50,000 | 167 | 298 |
| D2 | P7 | 15,923 | 15,897 | 15,923 | 175 | 90 |
| D3 | P7 | 5,071 | 5,124 | 5,071 | 168 | 30 |
| D4 | P7 | 1,615 | 1,658 | 1,615 | 169 | 9 |
| D5 | P7 | 514 | 517 | 514 | 153 | 3 |
| D6 | P7 | 164 | 212 | 164 | 135 | 1 |
| D7 | P7 | 52 | 109 | 52 | 99 | 0 |
| D8 | P7 | 17 | 52 | 17 | 55 | 0 |
| D9 | P7 | 5 | 10 | 5 | 26 | 0 |
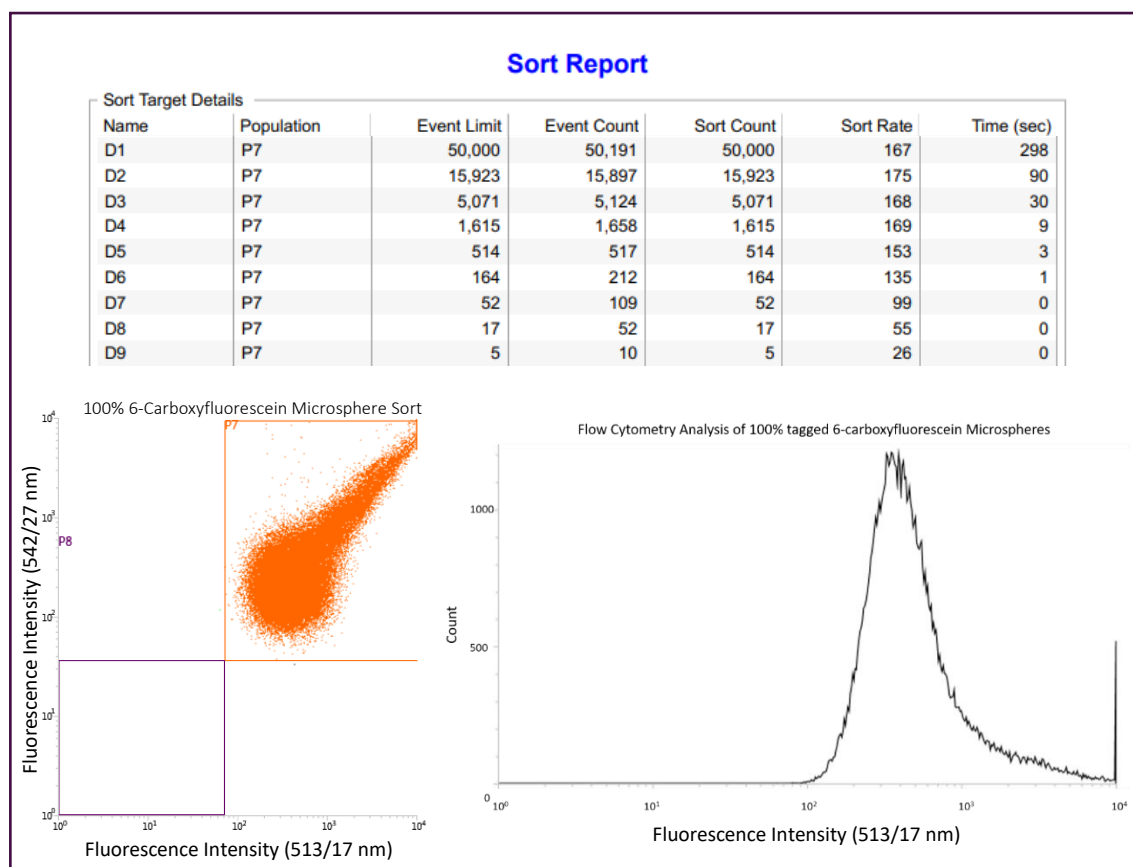
Figure 5.24. Flow Cytometry Analysis of 100% tagged 6-carboxyfluorescein microspheres sorting into a 96 well plate.

The sort ran successful, into the 96 well plate. The well plate was then screened on a plate reader to see if the level of fluorescence in the well was representative of the amount sorted into it. The results in figure 5.25 show the 96 well plate sort to be successful as it shows an exponential decay as the number of beads in the well becomes smaller.
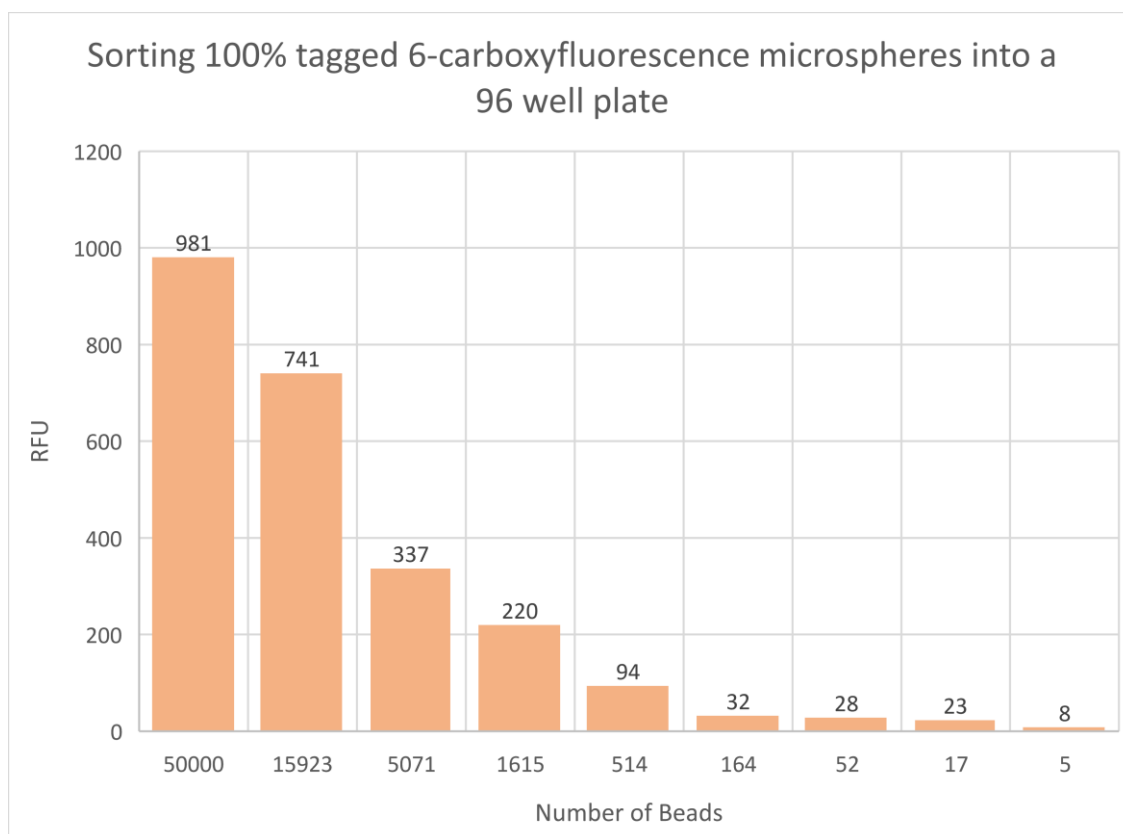
Figure 5.25. Plate Reader Analysis of 100% tagged 6-carbxyfluorescein microspheres sorted into a 96 well plate.

## 5.4.4. Sorting MinE07Lib to extract the top binding MinE07 modified aptamers.

The plan for sorting **MinE07Lib** was to do a two-way sort followed by a 96 well plate sort to pull out the top 96 binding **MinE07** aptamer microspheres. To target just the top 96 microspheres a limited amount of EGFR-Fc and Protein A-FITC was incubated with the **MinE07Lib**.

If each microsphere has the potential for the binding of $1.17e^{11}$ aptamer strands inside and out.

$$1.17e^{11} \times 96 = 1.1232e^{13}$$

Meaning that $1.1232e^{13}$ strands need to be targeted, and so this being the amount of EGFR molecules needed.

$$0.0001 \ g/mL \times 0.1 \ mL = 0.00001 \ g$$

$$0.0000\ 1g\ /\ 180000\ g/mol\ x\ 6.022e^{23}\ mol^{-1} = 3.345555556e^{13}\ EGFR\ molecules$$

Meaning that there are $3.345555556e^{13}$ EGFR molecules in 100 µL of a 0.1 mg/mL solution of the EGFR-Fc. This is more than is needed for the $1.1232e^{13}$ strands.

$$3.345555556e^{13}\ /\ 1.1232e^{13} = 2.979$$

$$0.1\ mg/mL\ /\ 2.979 = 0.03357\ mg/mL$$

The concentration to target just 96 microspheres is 0.03357 mg/mL. The original weight of microspheres put on the column for the synthesis of **MinE07Lib** was 0.0087 g which is ~13,107,200 microspheres, which allowed for 200 microsphere copies of each new modified **MinE07** aptamer. At the end of the library synthesis, the microspheres have moved location 88 times causing a large loss of microspheres. The final weight of the library being 0.0029 g equalling ~4,364,697.6 microspheres. This allows for 66 microsphere copies of each new modified **MinE07** aptamer. If all 66 were successfully synthesised for each aptamer, potentially only 2 different MinE07 aptamers would be in the top binders sorted, assuming sorting of 2 x 96 well plates of hits.

All of the **MinE07Lib** was incubated with EGFR-Fc followed by Protein A-FITC. The microspheres were spun and the bead pellet re-suspended in sheath fluid and added to a FACS tube. The FACSJazz™ lasers were calibrated before the sample was run with the 8 peak BD beads. The FACS two-way sort was calibrated using the Accudrop set up program along with the Accudrop beads. The plain microspheres were first run on the FACS to optimise the conditions and gating.
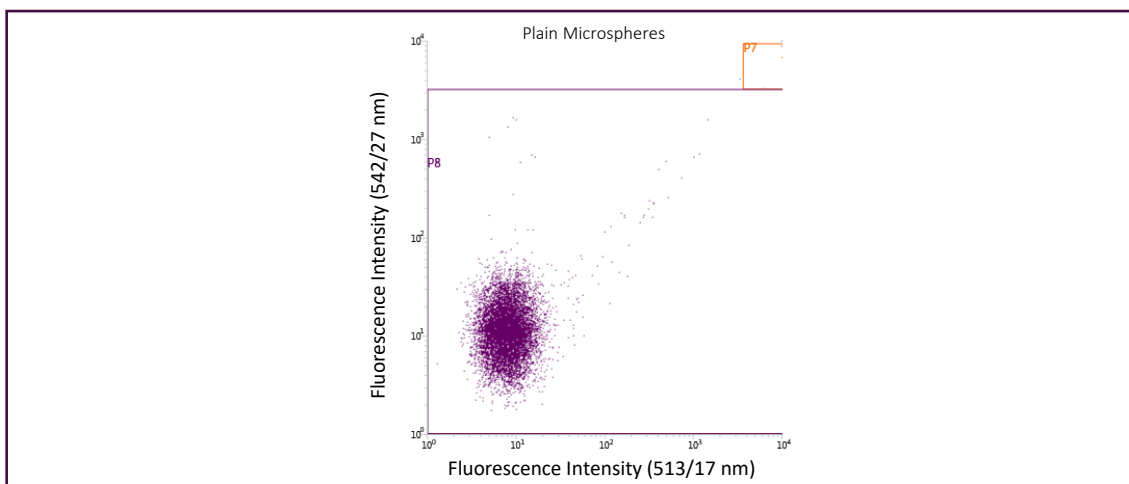
Figure 5.26. FACS data of the plain microspheres to select gating for sorting.

Gates orange P7 and purple P8 were set up from the plain microsphere data to allow for sorting of the top binding MinE07 aptamers. Round 1 of sorting sorted a total of 5,648,248 microspheres of which 23,176 were sorted into gate orange P7 (figure 5.27).
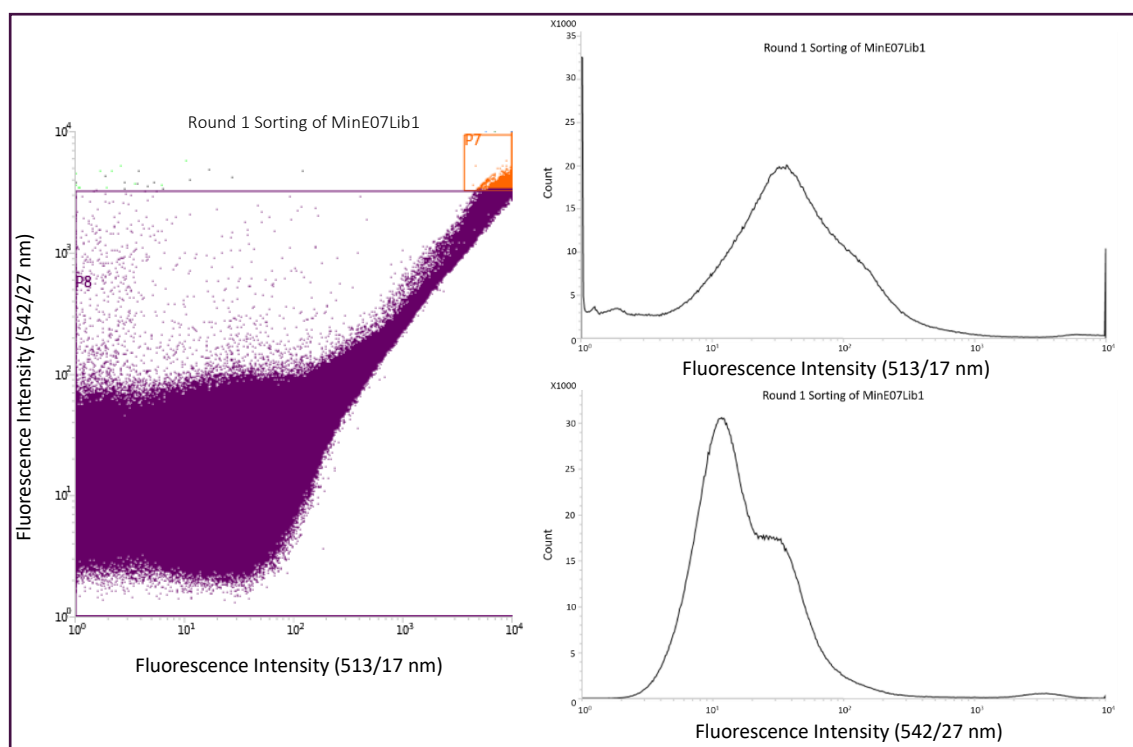


Figure 5.27. FACS data of round 1 of sorting MinE07Lib.

The microspheres sorted into purple P8 gate were kept separate and no longer sorted. The 23,176 microspheres that were sorted into gate orange P7 were then carried onto the next

round of sorting. The data from the first sort compared with other samples can be found in figure 5.28. The bottom histogram shows the **MinE07Lib** compared with the plain microspheres, the MinE07ARP-EGFR-Fc-Protein A-FITC Microspheres and the 100% tagged 6-carboxyfluorescein microspheres.
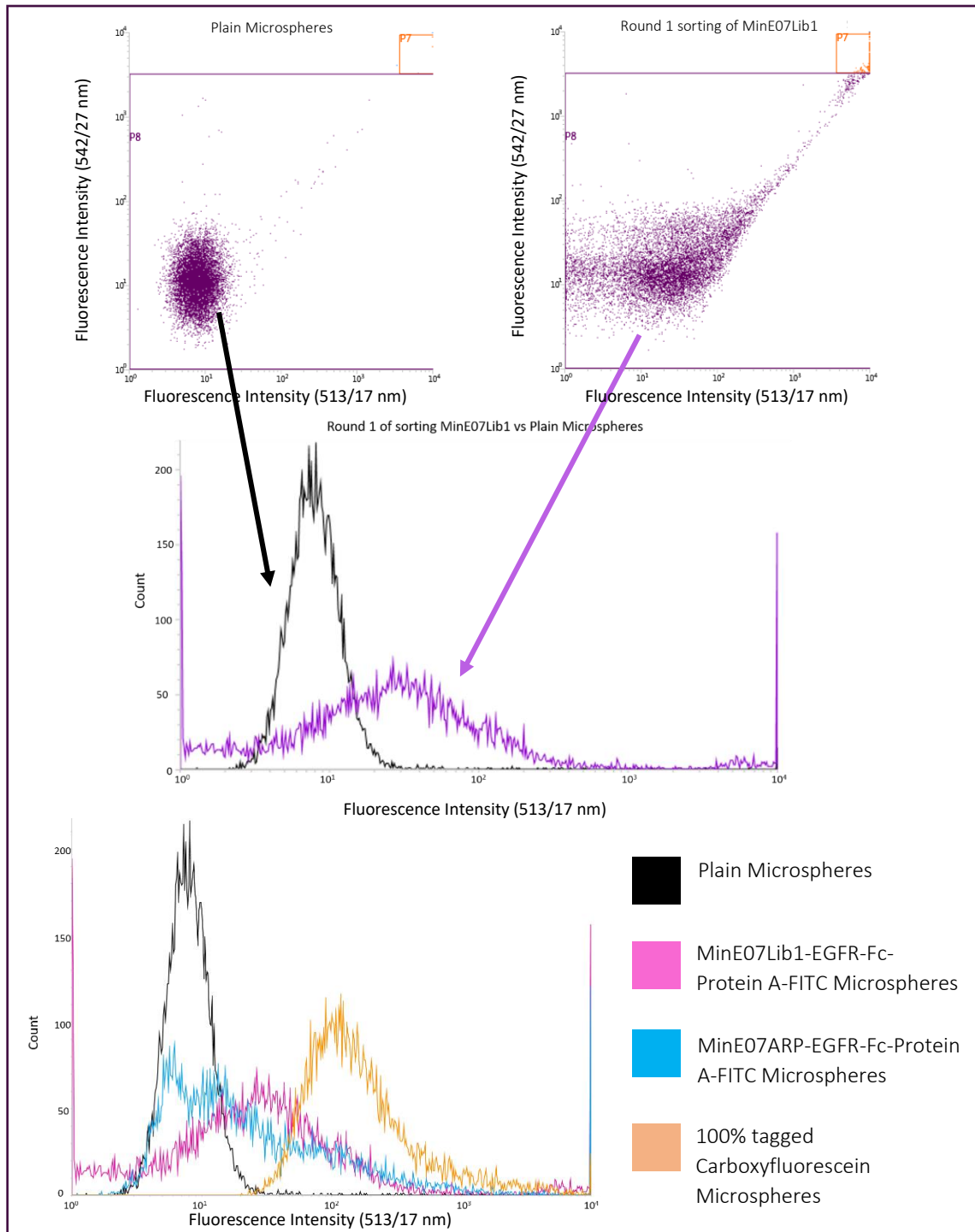


Figure 5.28. FACS data of the first sort of MinE07Lib, conditions optimised by with plain microspheres.

Round 2, sorted 4,627 microspheres into orange P7 gate out of the 23,176 from the previous round. The orange P7 gate was slightly adjusted to allow for some photo-bleaching of the samples (figure 5.29).
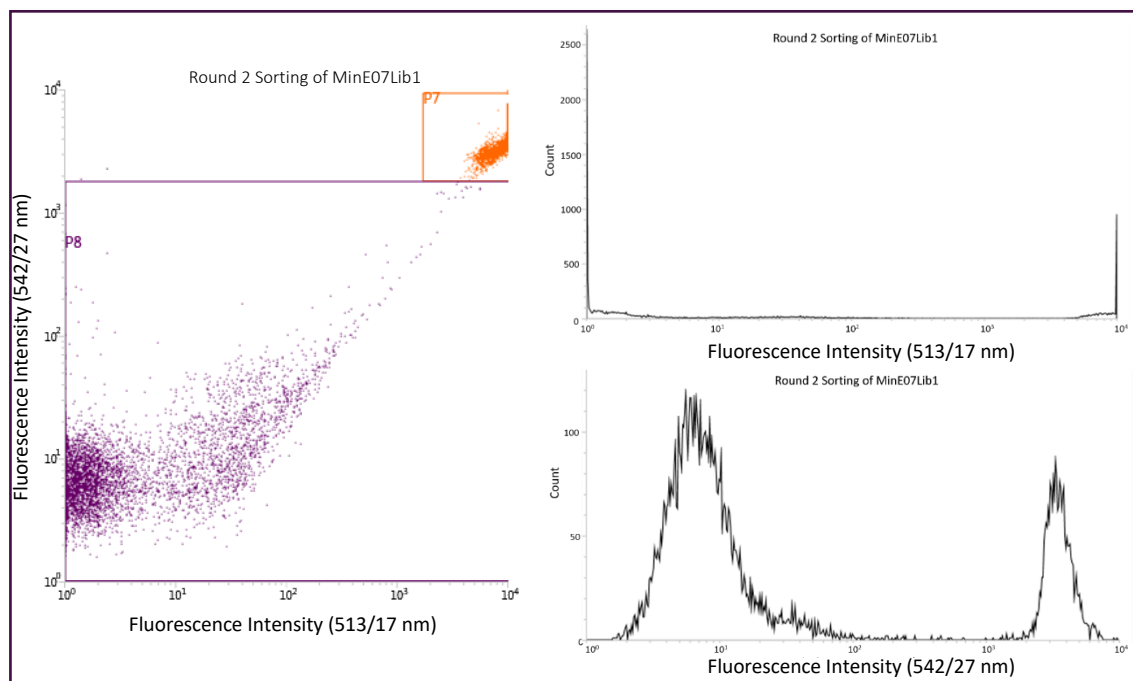


Figure 5.29. FACS data of round 2 of sorting MinE07Lib.

4,627 microspheres is more than the desired 96 microspheres so a 3rd round of sorting was performed. Round 3 sorted 339 microspheres into orange P7 gate out of the 4,627 from the previous round. The orange P7 gate was slightly adjusted to allow for some photo-bleaching of the samples (figure 5.30).
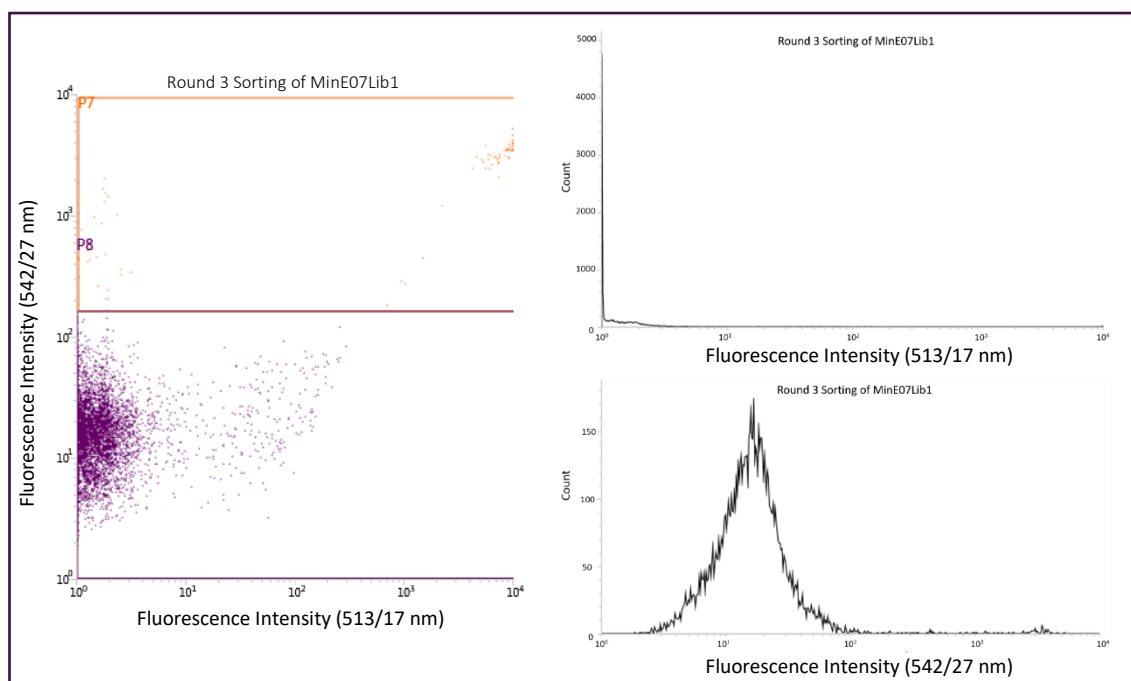
Figure 5.30. FACS data of round 3 of sorting MinE07Lib.

339 is low enough to start sorting into 96 well plates. The FACS 96 well plate sort was calibrated using the 96 well plate program set up using the accudrop beads. Once the stream was calibrated to sort into the well of the 96 wells plate, a fresh well plate was used and 100 µL of Milli Q® water was added to each well. This stops the droplet that is going into that well from bouncing back out. From the orange P7 gate, 96 microspheres were sorted individually into the well of the first 96 well plate. In the second well plate 74 microspheres were sorted. Giving a total of 170 microspheres that were sorted as the top binding aptamers. Anything from purple P8 gate was not sorted and discarded (figure 5.31).
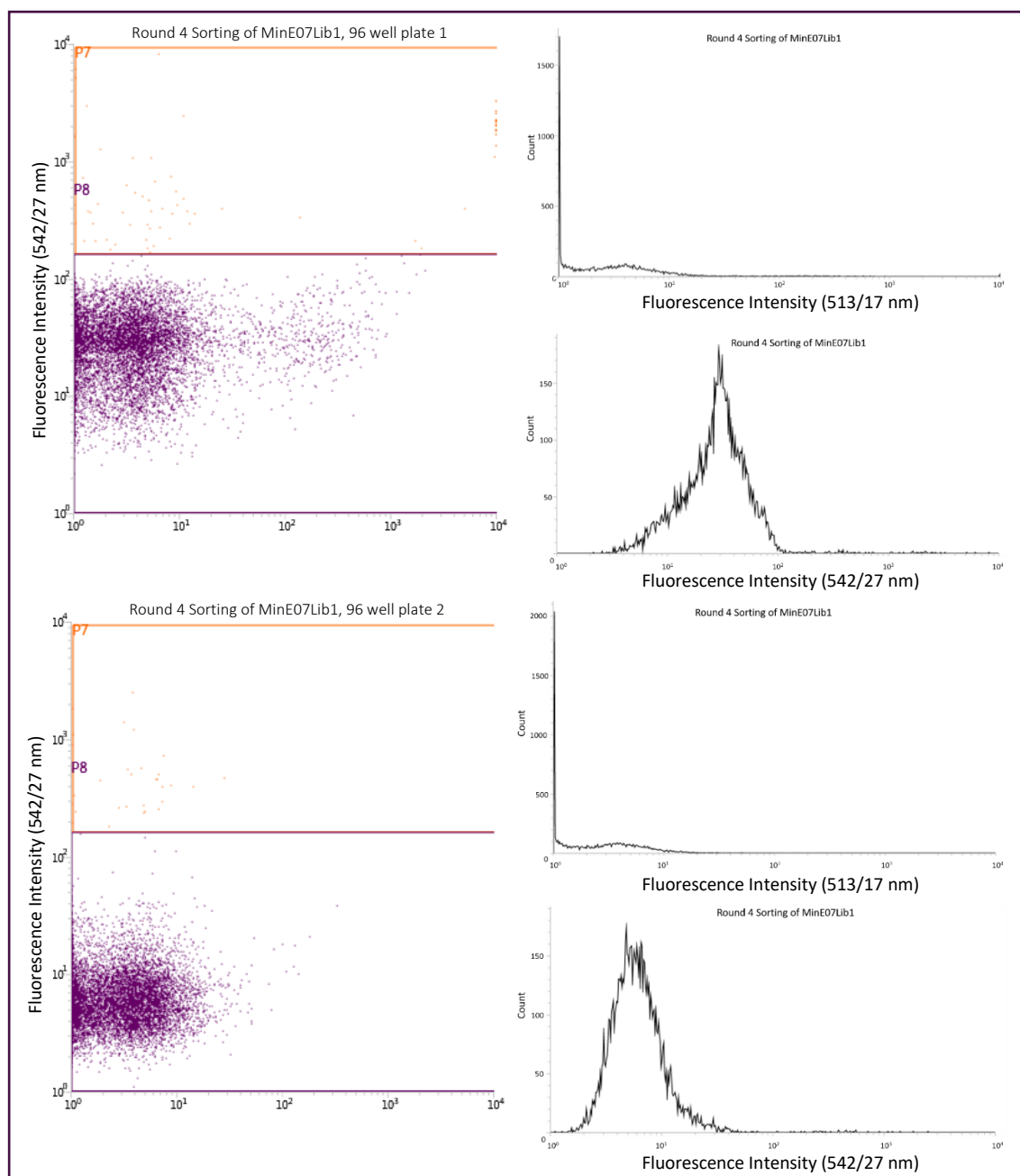
Figure 5.31. FACS data of round 4 of sorting MinE07Lib into 96 well plates.

## 5.4.5. Photo-cleaving the top aptamers from the microspheres

Once the top 170 microspheres were sorted in to 2x96-well plates, the aptamers were photocleaved from the microspheres into the Milli Q® water. This photo-cleavable linker was the first molecule bound to the microspheres. It can be cleaved using long wave UV light in the range of 340-365 nm[459,460,461,462]. The photocleaving mechanism is shown in figure 5.32. The

samples were left in the 96-well plates, longwave UV light was irradiated from 15-20 cm above[463] using a UV lamp for 3 hours. After irradiation, the solutions from the well were collected.
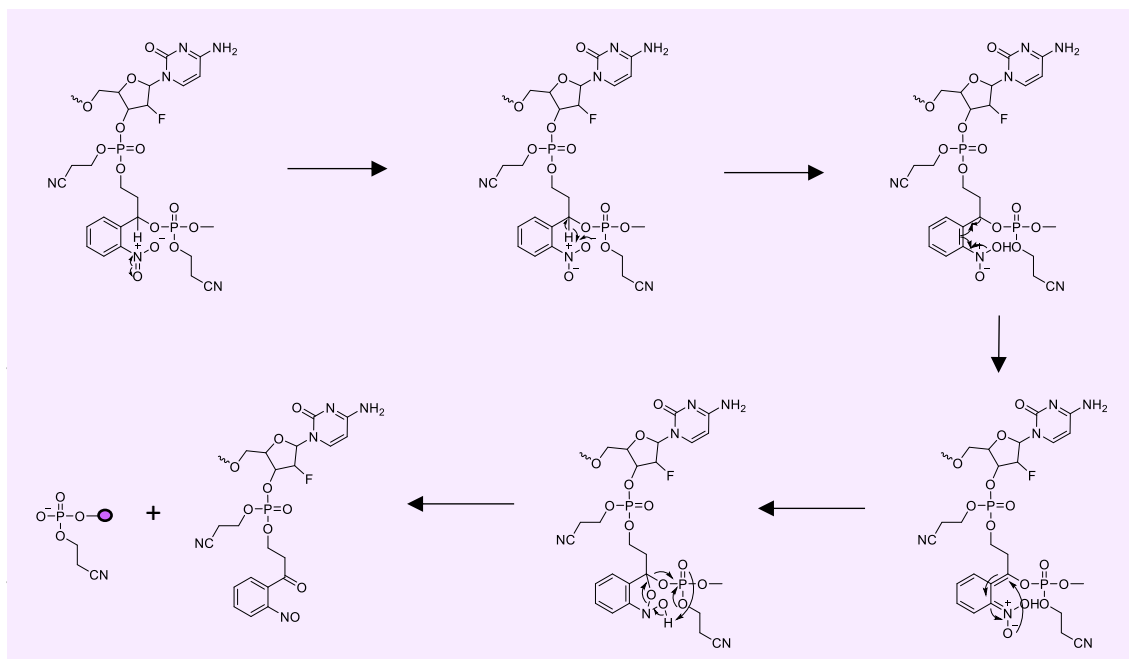


Figure 5.32. Mechanism of the photocleaving from the TentaGel microspheres.

## 5.4.6. Tandem Mass Spectrometry of MinE07-Biotin

### 5.4.6.1. LC-MS/MS Analysis of MinE07-Biotin

Comprehensive characterisation of nucleotide sequences of oligonucleotide can be achieved by reverse-phase liquid chromatography coupled with mass spectrometry. For separation of oligonucleotide mixture using a C18 reverse-phase column, the negative charge on the phosphate backbone must be masked by tetraethylammonium bromide. As well as acting as an ion pairing reagent, this reagent makes the oligonucleotides more hydrophobic and allows chromatographic retention. For this reason mobile phase A was chosen to be 8 mM tetraethyl ammonium bromide (TEAB) in LC-MS grade water (pH 7.5-7.8) and mobile phase B was 8 mM TEAB in a 1:1 ratio of LC-MS grade methanol and water. Aptamer MinE07-Biotin was run at 10 µM for the initial run. MinE07- Biotin aptamer was run first in order to establish a method for when running the top selected aptamers from the library.
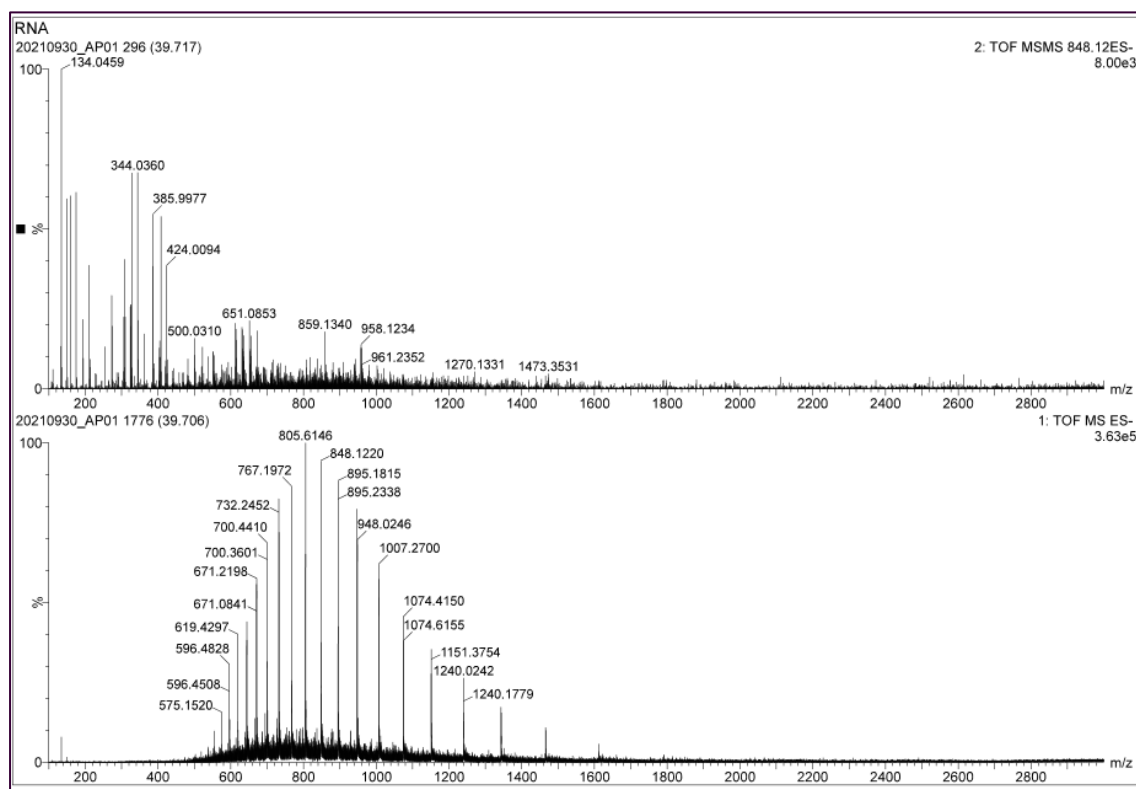


Figure 5.33. LC-MS/MS spectra of MinE07-Biotin at 10 µM.

Figure 5.33 shows the MS and MS/MS spectra of **MinE07-Biotin** (10 µM) at retention time (RT) 39.72 minutes. The data shows the correct molecular weight of 16093 Da. As this method was successful, aptamer MinE07-Biotin was then run at 1 µM, 0.5 µM, 0.2 µM, 0.1 µM, 0.05 µM, 0.02 µM and 0.01 µM to work out limit of detection (LOD).
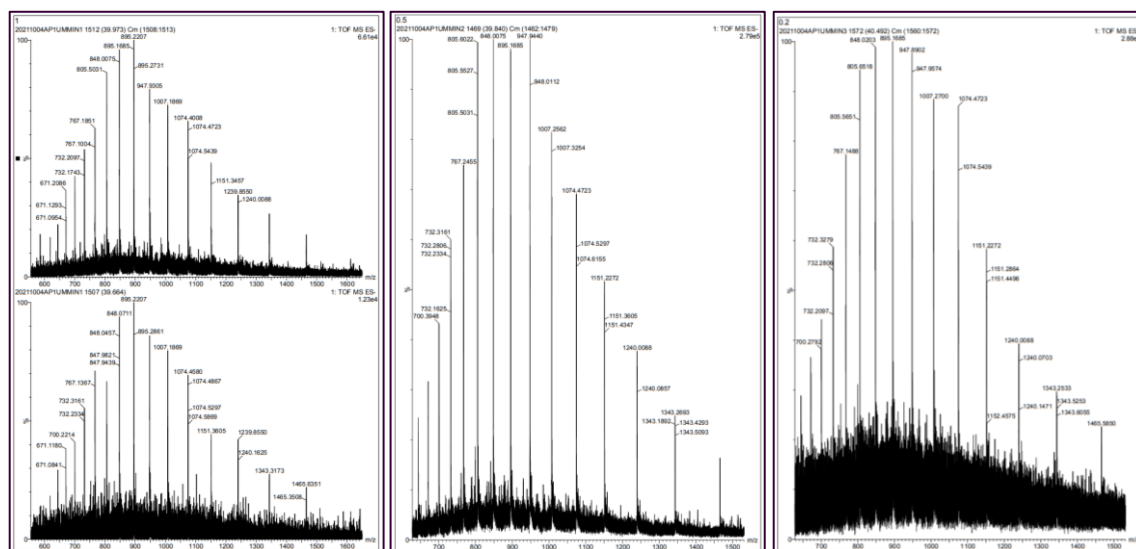


Figure 5.34. LC-MS Spectra of MinE07-Biotin at 1 µM. 0.5 µM and 0.2 µM.

Figure 5.34 shows the MS data of **MinE07-Biotin** at the concentrations of 1, 0.5 and 0.2 µM. The signal to noise ratio decreases as the concentration decreases.
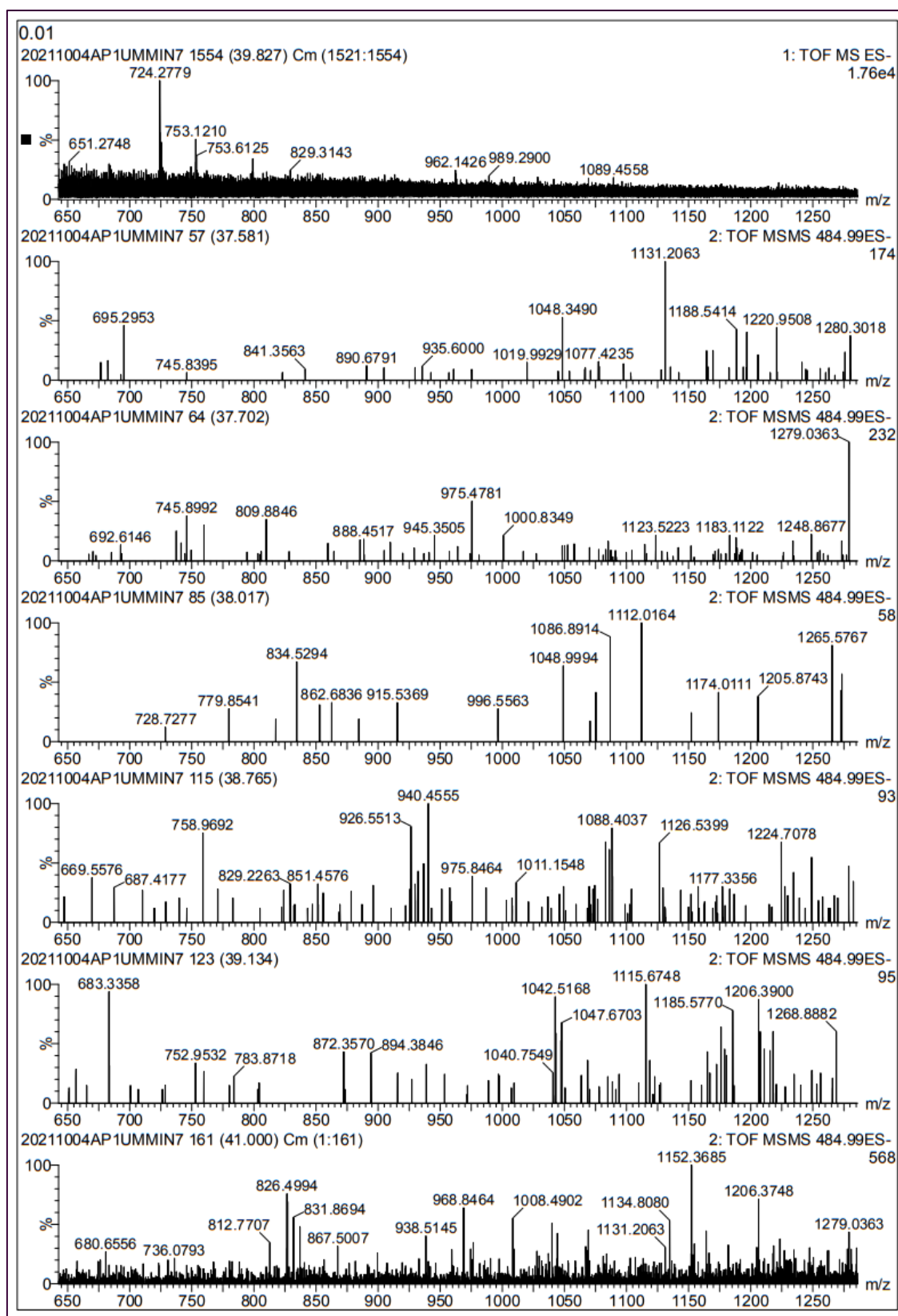
Figure 5.35. LC-MS/MS Fragment spectra of MinE07-Biotin at 0.01 μM.

Figure 5.35 shows the LC-MS and LC-MS/MS fragment data of **MinE07-Biotin** at the concentrations of 0.01 μM. The data shows that this the limit of detection (LOD) for LC-MS/MS.

The **MinE07Lib** was synthesised at a concentration of 1 µmol, however this may not be the concentration after sorting, photocleaving and being moved from tube, so it is important that the LOD of **MinE07** be as low as possible. It could potentially be a 10th of the initial concentration.

### 5.4.6.2. Fragment Pattern Data of MinE07-Biotin

MinE07-Biotin: 5'-Biotin-rGrGrAfCrGrGrAfUfUfUrArAfUfCrGfCfCrGfUrArGrArArArArGfCrAfUrGf UfCrArArArGfCfCrGrGrArAfCfCrGfUfCfC-3'
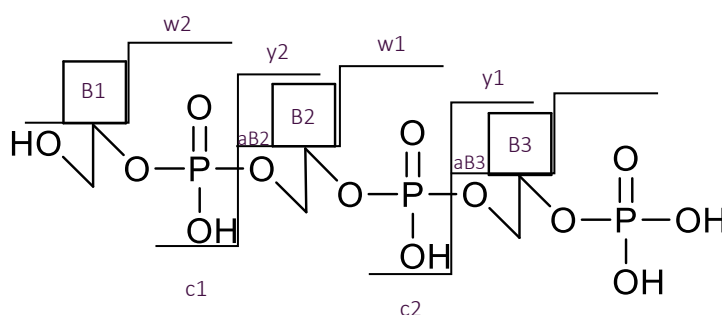


Figure 5.36. Oligonucleotides Fragment Pattern

The most abundant fragment ions are c- and y-. w- and ab- are also often observed (figure 5.36). These are produced by certain breaks in the oligonucleotide chain caused by the collision-induced dissociation. In the 5' direction the c- notation is the nucleotide molecular weight and the -ab(n) notation is c(n-1) + (B(n) − Base MW). In the 3' direction y notation is the nucleotide molecular weight and w(n) = y(n) + $PO_3H$ molecular weight (80 Da). The fragment pattern of **MinE07-Biotin** was determined before the **MinE07Lib** hits were run so that the fragment patterns of the hits could be compared to it. Figure 5.37 shows all the discovered fragments in **MinE07-Biotin.** This figure shows all the overlapping fragments, the more overlap the better, as you can then be more sure of the data identifying modifications. All individual fragment data are shown below in figures 5.38-5.59.

5'-Biotin- **| rGrGrAfCrGrG |** rAfUfUfUrA **| rAfUfCrGfCfC |** rGfUrArGrArArA **|** rA **| rGfCrAfUrG |** fUfCrA **| rArArGfCfCrG |** rGrArA **| fCfCrGfUfCfC**-3'

**| GrGrAfUfUfU |**                                         **| rArArArArGfC |**         **| rGfUfCrArArArG | | fCfCrGrGrArA |**

                                                                                                                  **| rGrGrArAfCfC |**

Figure 5.37. MinE07-Biotin Fragment Pattern. Fragments are shown by the bold sections.

### 5.4.6.2.1. Fragment 1 of MinE07-Biotin

5'-Biotin- **| rGrGrAfCrGrG |** rAfUfUfUrArAfUfCrGfCfCrGfUrArGrArArArArGfCrAfUrGfUfCrArArA

rGfCfCrGrGrArAfCfCrGfUfCfC-3'
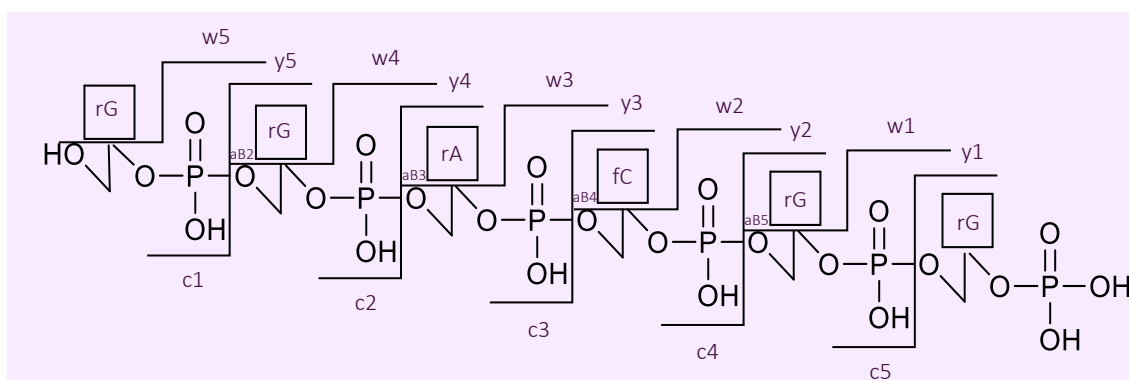


Figure 5.38. Fragment 1 of MinE07-Biotin 5'-3'.
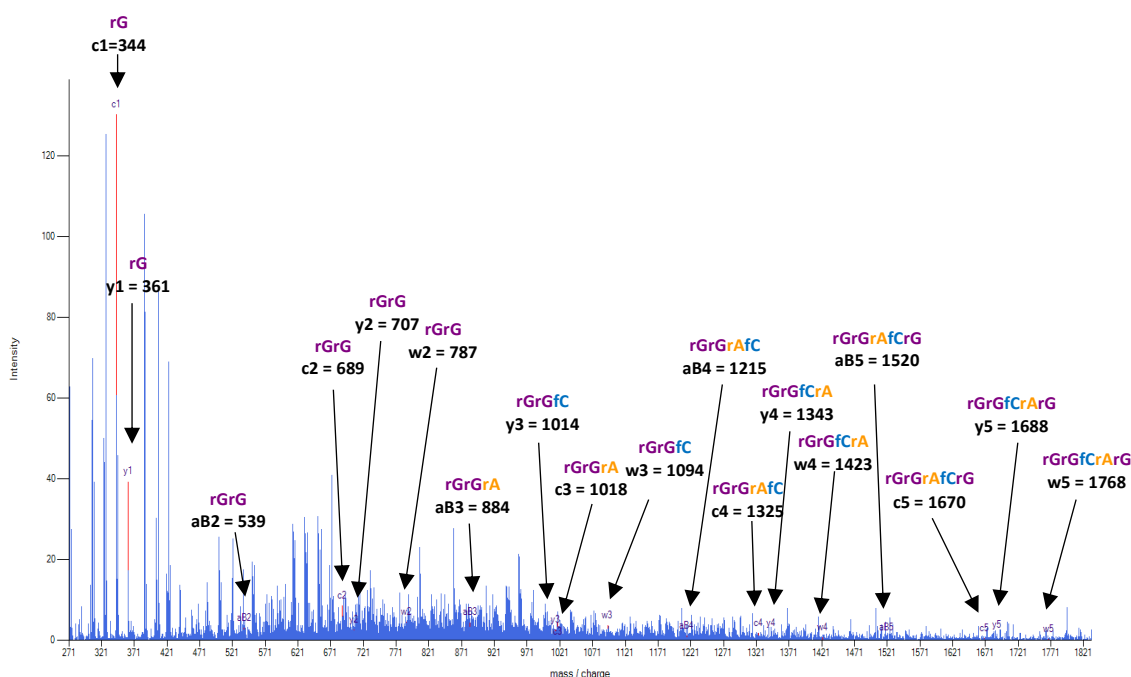


Figure 5.39. LC-MS/MS Fragment 1 Data of MinE07-Biotin.

5'-Biotin-rGrGrAfCr **| GrGrAfUfUfU |** rArAfUfCrGfCfCrGfUrArGrArArArArGfCrAfUrGfUfCrArArA

rGfCfCrGrGrArAfCfCrGfUfCfC-3'



Figure 5.40. Fragment 2 of MinE07-Biotin 5'-3'.
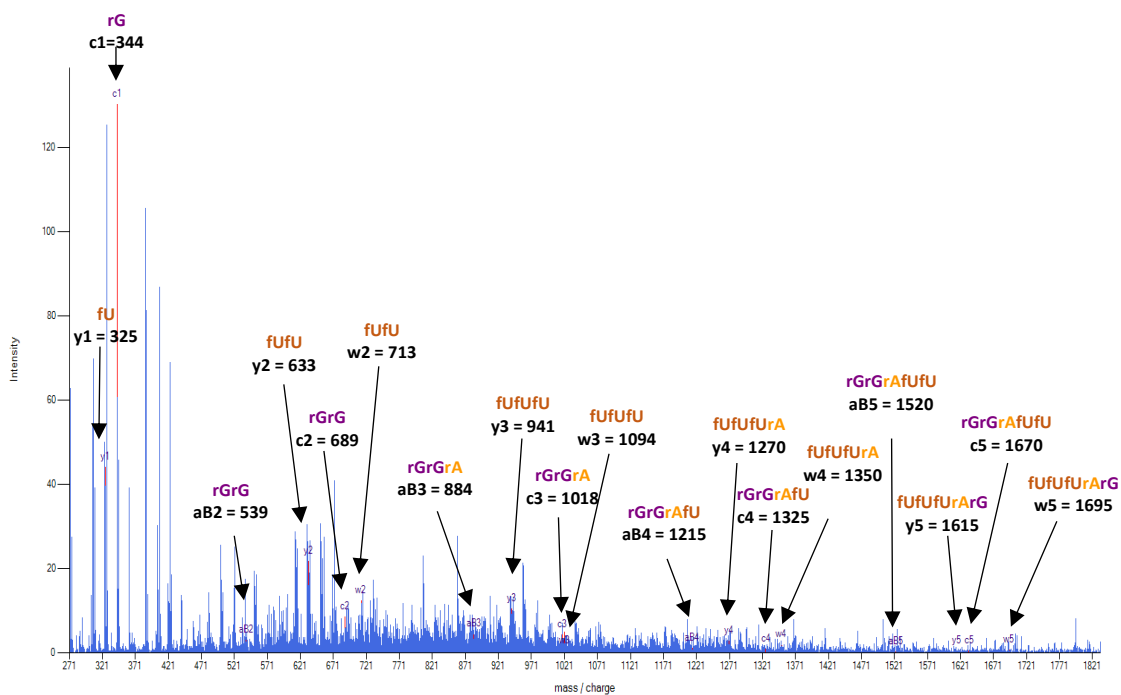


Figure 5.41. LC-MS/MS Fragment 2 Data of MinE07-Biotin.

5'-Biotin-rGrGrAfCrGrGrAfUfUfUrA **| rAfUfCrGfCfC |** rGfUrArGrArArArArGfCrAfUrGfUfCrArArA

rGfCfCrGrGrArAfCfCrGfUfCfC-3'



Figure 5.42. Fragment 3 of MinE07-Biotin 5'-3'.
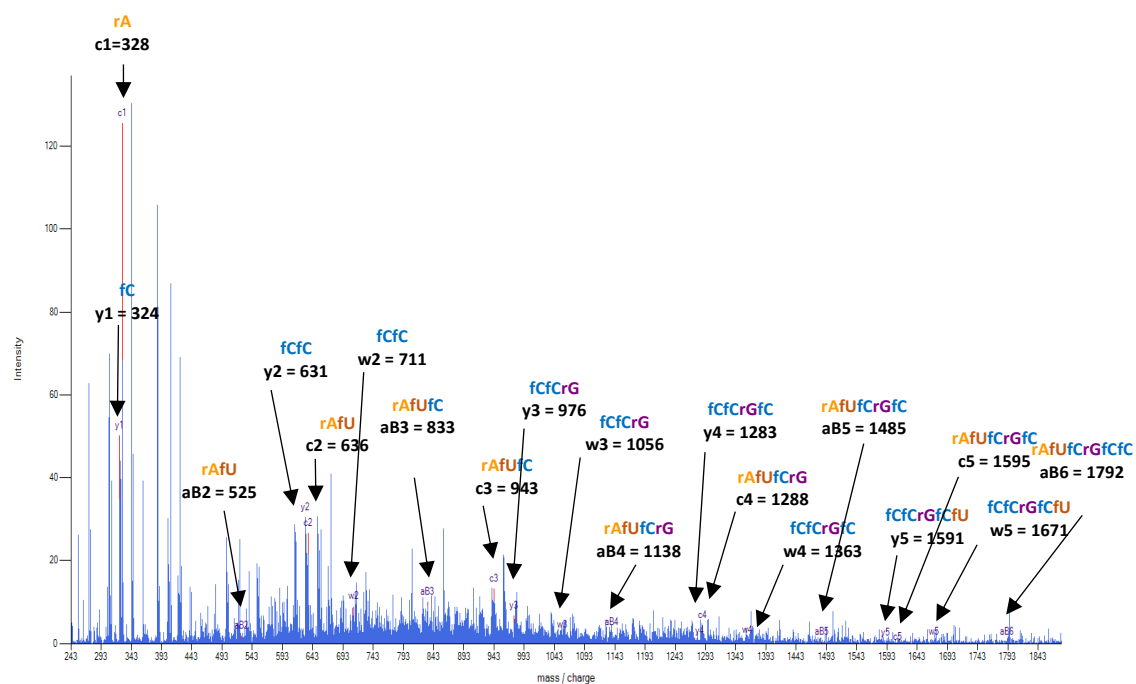


Figure 5.43. LC-MS/MS Fragment 3 Data of MinE07-Biotin.

### 5.4.6.2.4. Fragment 4 of MinE07-Biotin

5'-Biotin-rGrGrAfCrGrGrAfUfUfUrArAfUfCrGfCfC **| rGfUrArGrArArA |** rArGfCrAfUrGfUfCrArArA

rGfCfCrGrGrArAfCfCrGfUfCfC-3'
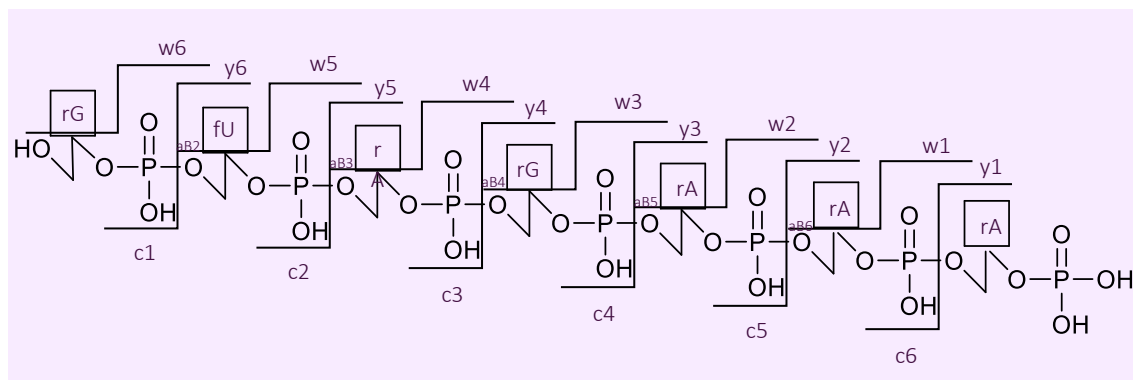


Figure 5.44. Fragment 4 of MinE07-Biotin 5'-3'.
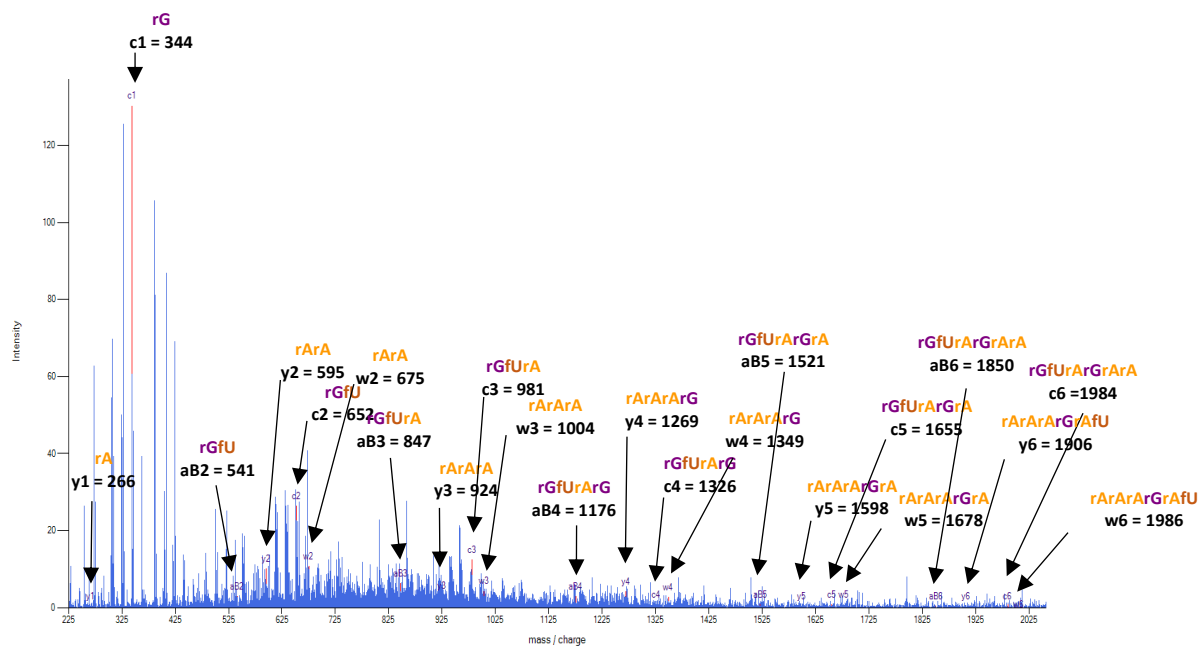


Figure 5.45. LC-MS/MS Fragment 4 Data of MinE07-Biotin.

### 5.4.6.2.5. Fragment 5 of MinE07-Biotin

5'-Biotin-rGrGrAfCrGrGrAfUfUfUrArAfUfCrGfCfCrGfUrArG **| rArArArArGfC |** rAfUrGfUfCrArArA

rGfCfCrGrGrArAfCfCrGfUfCfC-3'
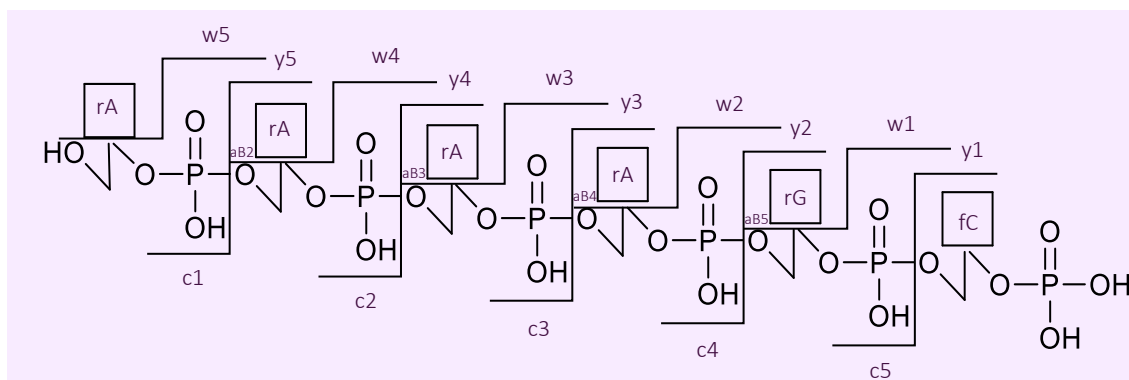


Figure 5.46. Fragment 5 of MinE07-Biotin 5'-3'.
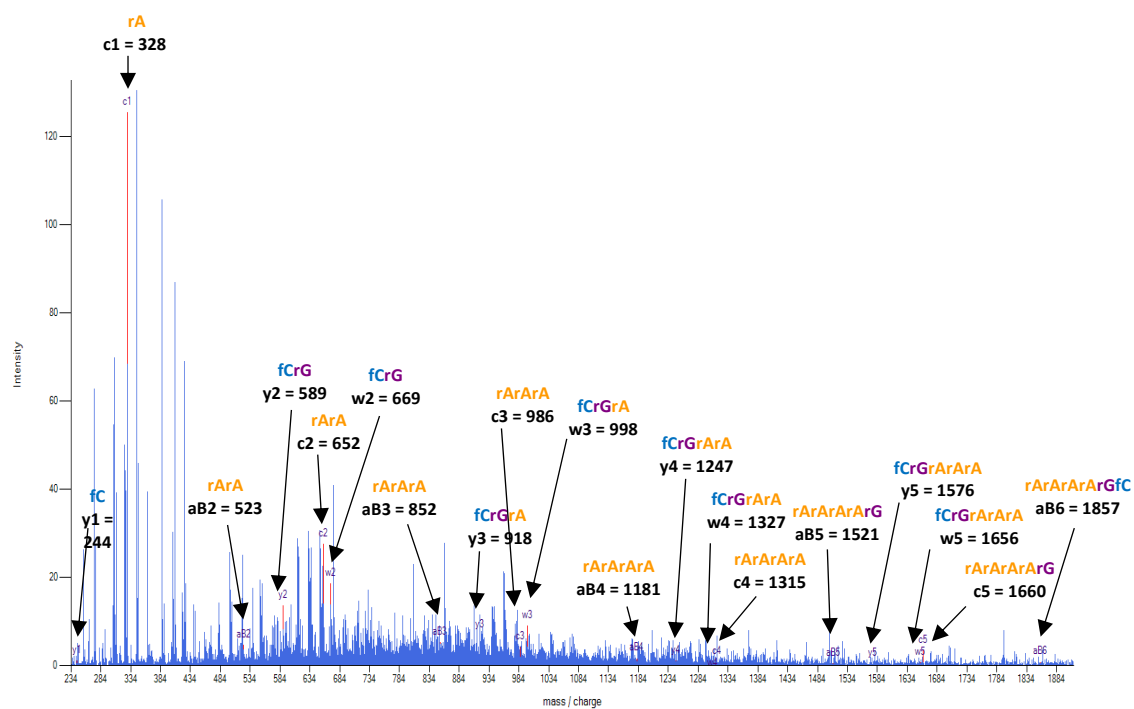


Figure 5.47. LC-MS/MS Fragment 5 Data of MinE07-Biotin.

## 5.4.6.2.6. Fragment 6 of MinE07-Biotin

5'-Biotin-rGrGrAfCrGrGrAfUfUfUrArAfUfCrGfCfCrGfUrArGrArArArA **| rGfCrAfUrG |** fUfCrArArA

rGfCfCrGrGrArAfCfCrGfUfCfC-3'



Figure 5.48. Fragment 6 of MinE07-Biotin 5'-3'.
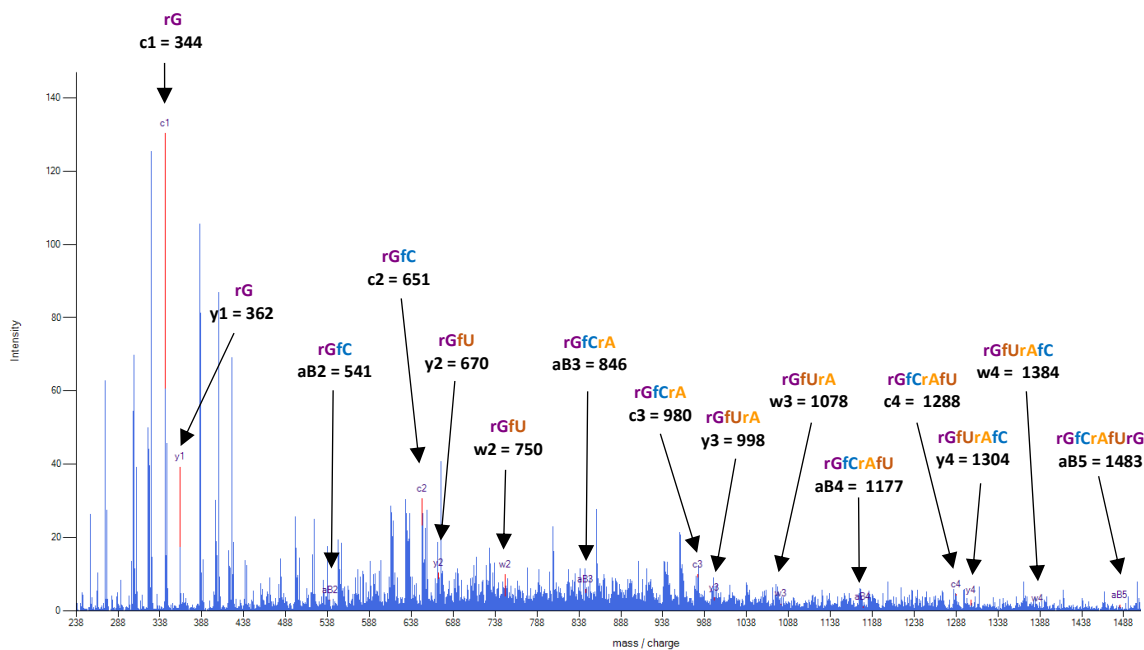


Figure 5.49. LC-MS/MS Fragment 6 Data of MinE07-Biotin.

### 5.4.6.2.7. Fragment 7 of MinE07-Biotin

5'-Biotin-rGrGrAfCrGrGrAfUfUfUrArAfUfCrGfCfCrGfUrArGrArArArArGfCrAfU **| rGfUfCrArArArG**

**|** fCfCrGrGrArAfCfCrGfUfCfC-3'
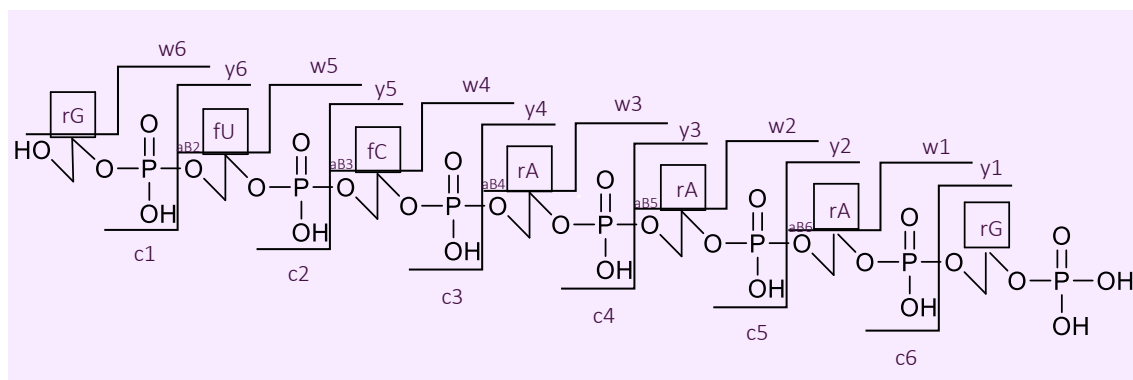


Figure 5.50. Fragment 7 of MinE07-Biotin 5'-3'.
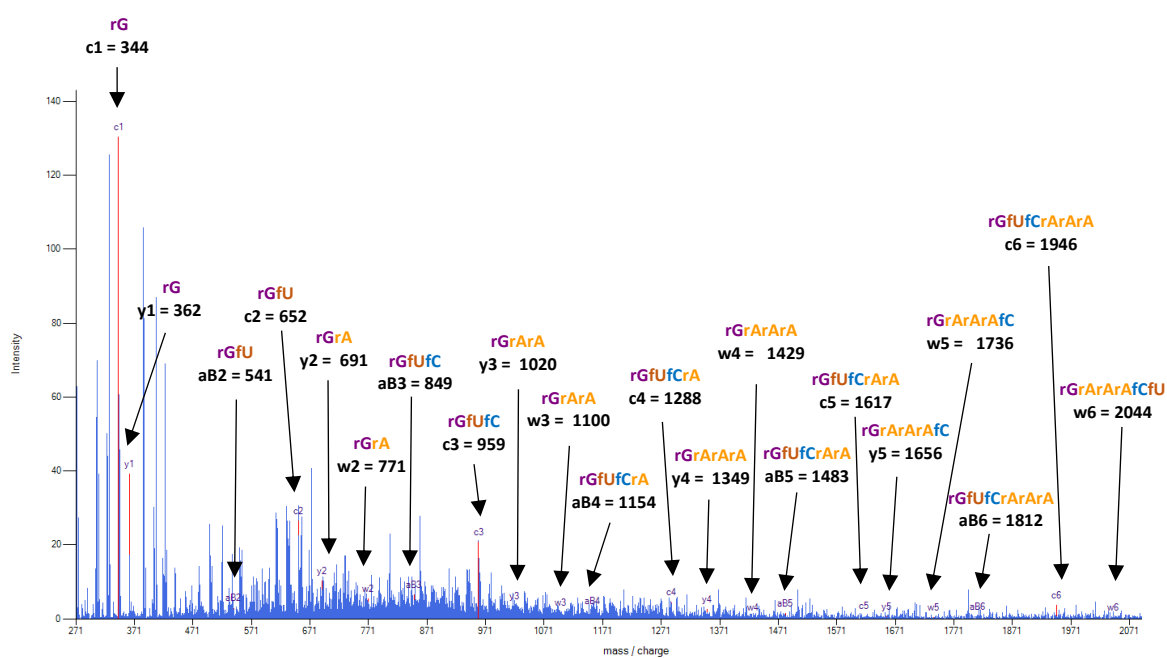


Figure 5.51. LC-MS/MS Fragment 7 Data of MinE07-Biotin.

### 5.4.6.2.8. Fragment 8 of MinE07-Biotin

5'-Biotin-rGrGrAfCrGrGrAfUfUfUrArAfUfCrGfCfCrGfUrArGrArArArArGfCrAfUrGfUfCrA **| rArArG**

**fCfCrG |** rGrArAfCfCrGfUfCfC-3'
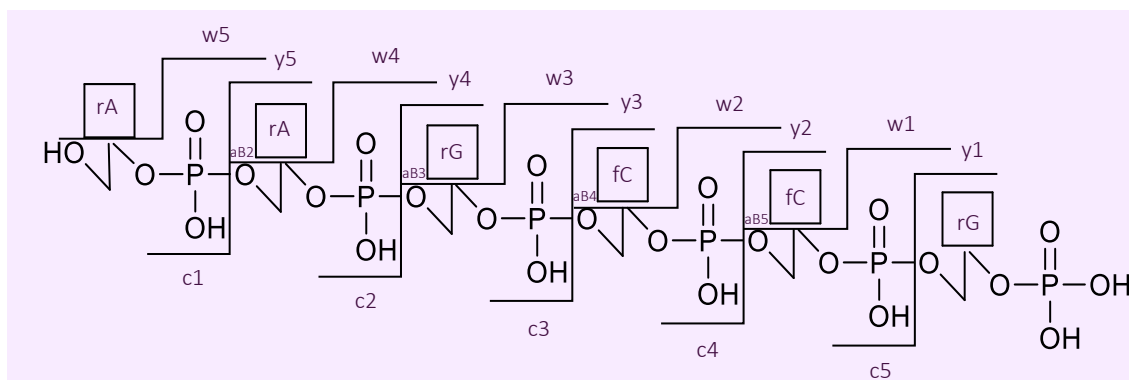


Figure 5.52. Fragment 8 of MinE07-Biotin 5'-3'.
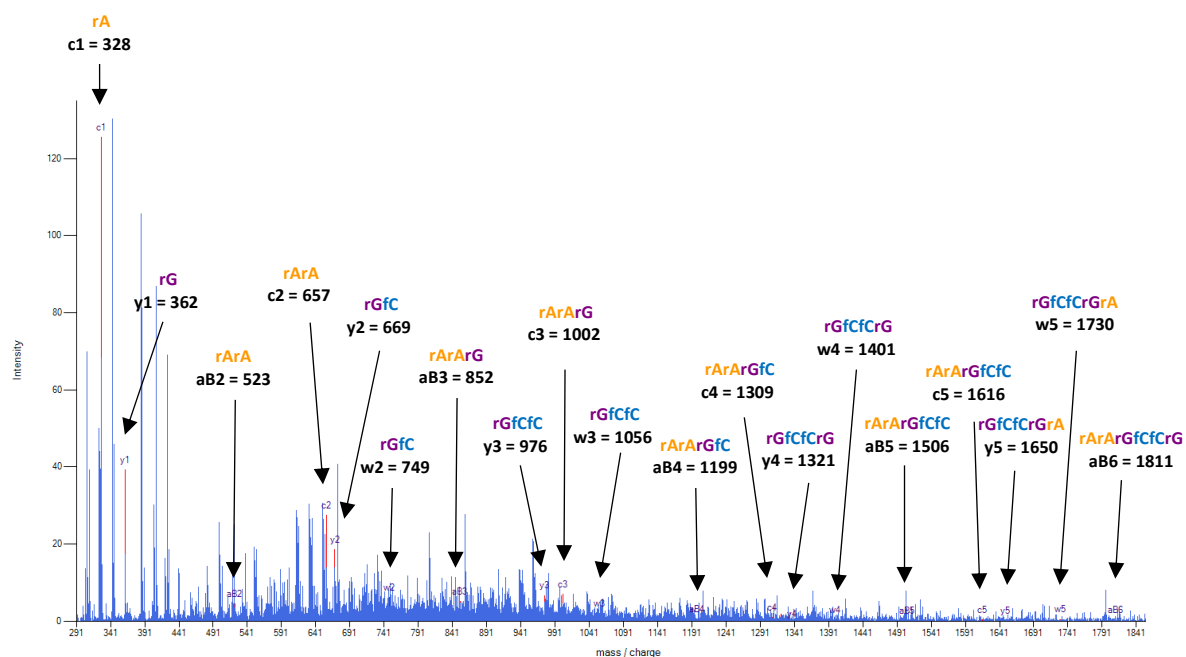


Figure 5.53. LC-MS/MS Fragment 8 Data of MinE07-Biotin.

### 5.4.6.2.9. Fragment 9 of MinE07-Biotin

5'-Biotin-rGrGrAfCrGrGrAfUfUfUrArAfUfCrGfCfCrGfUrArGrArArArArGfCrAfUrGfUfCrArArArG

**| fCfCrGrGrArA |** fCfCrGfUfCfC-3'
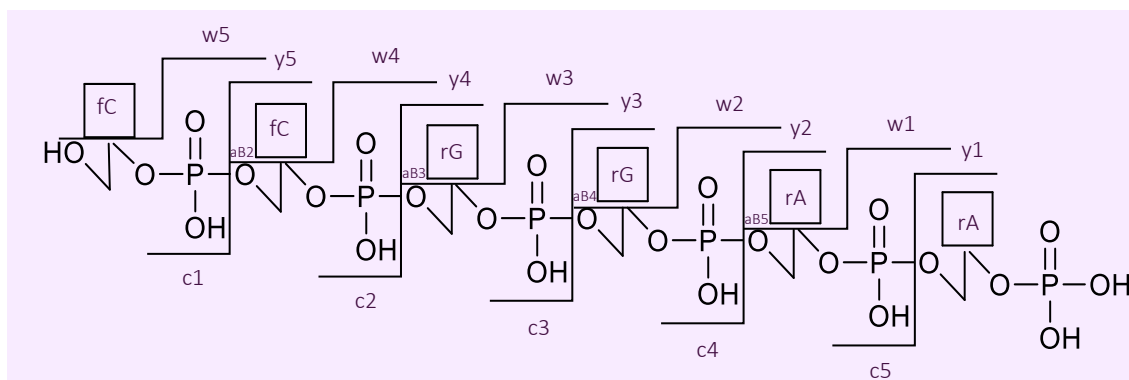


Figure 5.54. Fragment 9 of MinE07-Biotin 5'-3'.
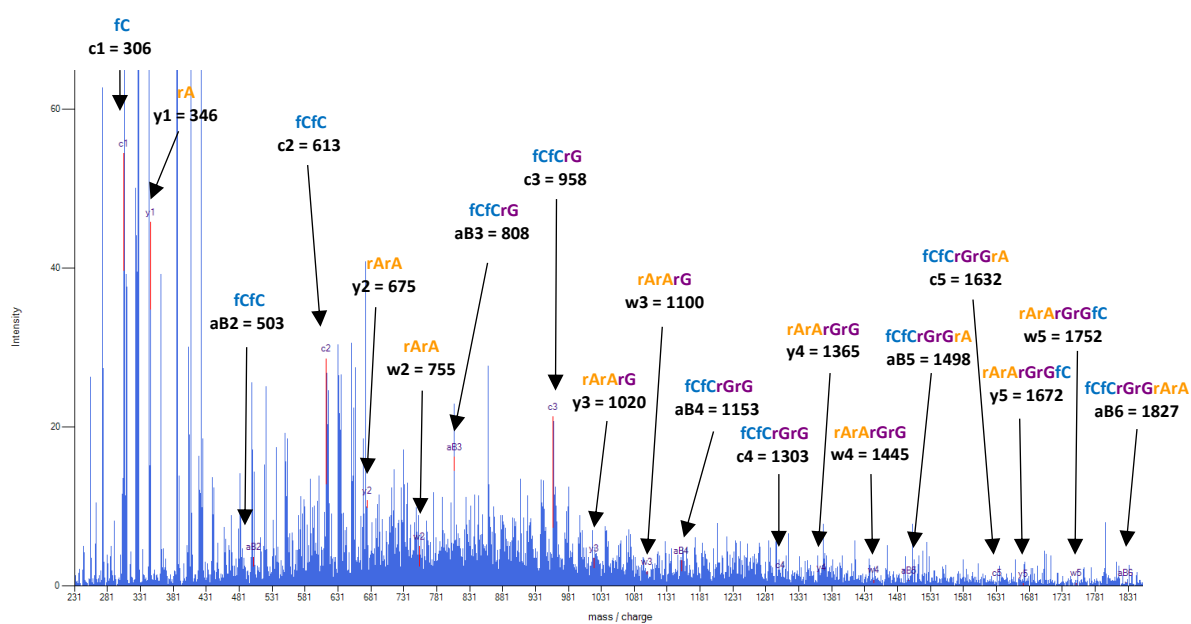


Figure 5.55. LC-MS/MS Fragment 9 Data of MinE07-Biotin.

5'-Biotin-rGrGrAfCrGrGrAfUfUfUrArAfUfCrGfCfCrGfUrArGrArArArArGfCrAfUrGfUfCrArArArGfCf

C **| rGrGrArAfCfC |** rGfUfCfC-3'



Figure 5.56. Fragment 10 of MinE07-Biotin 5'-3'.
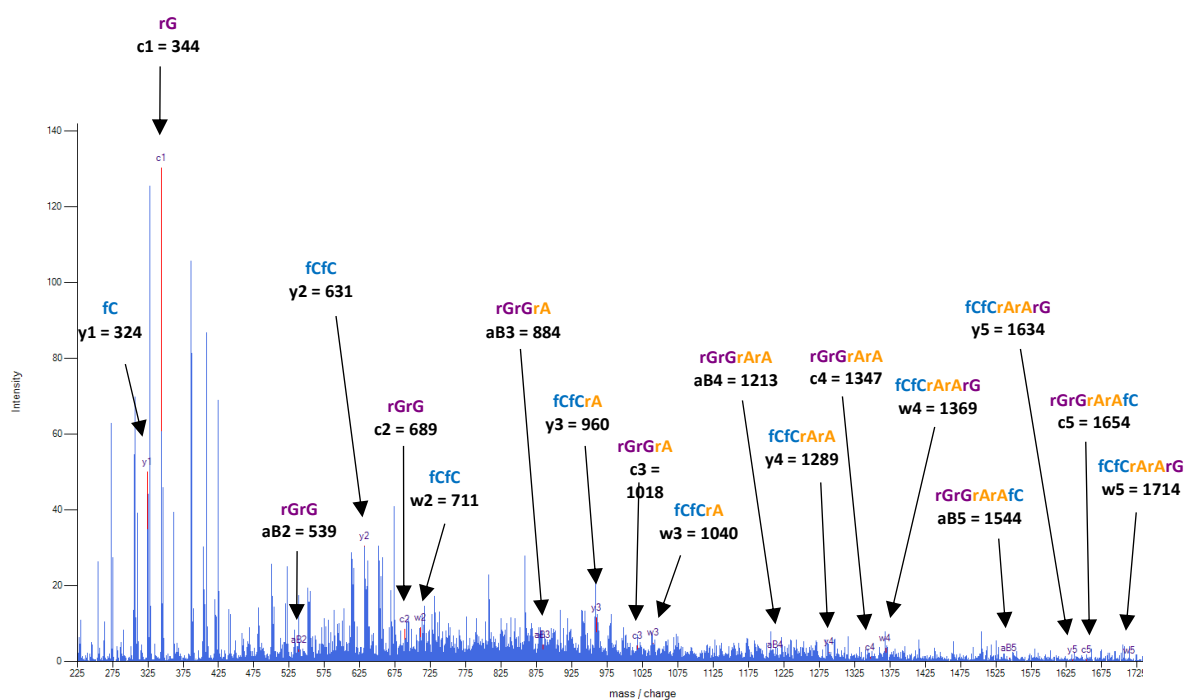


Figure 5.57. LC-MS/MS Fragment 10 Data of MinE07-Biotin.

300

5'-Biotin-rGrGrAfCrGrGrAfUfUfUrArAfUfCrGfCfCrGfUrArGrArArArArGfCrAfUrGfUfCrArArArGfCf

CrGrGrArA **| fCfCrGfUfCfC |**-3'



Figure 5.58. Fragment 11 of MinE07-Biotin 5'-3'.
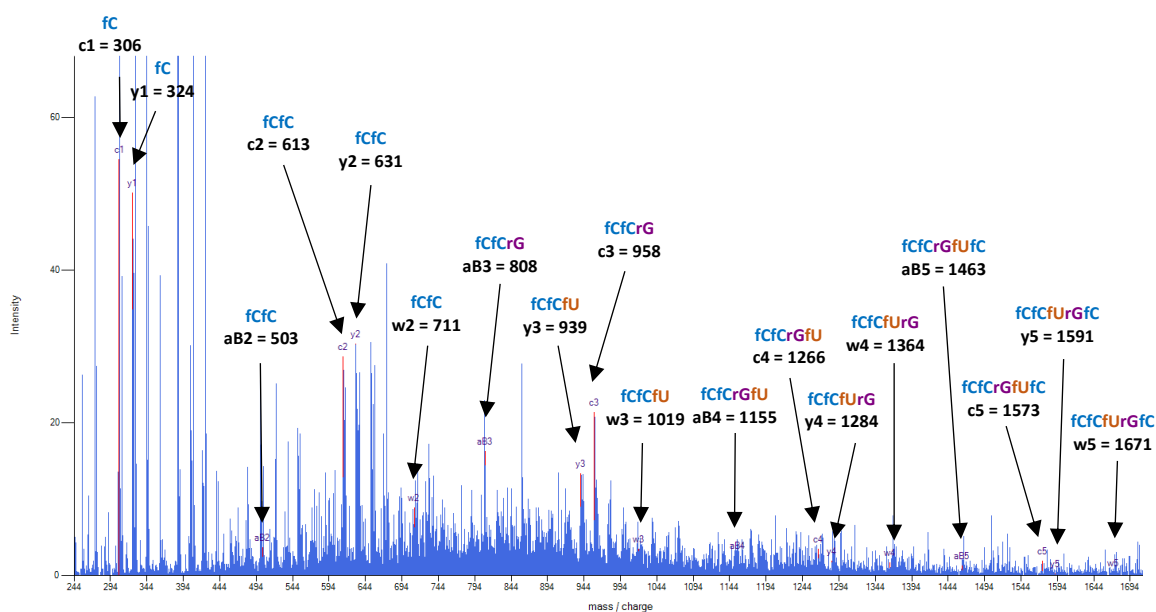


Figure 5.59. LC-MS/MS Fragment 11 Data of MinE07-Biotin.

This fragment data of **MinE07** shows that this MS/MS method works for aptamers of this length.
It breaks the aptamer into 11 fragments each of 4-7 nucleoside length, which makes the data
simple to locate within the correct section of **MinE07**. The fragment patterns of the top hit
aptamers were then compared to this pattern.

## 5.4.7. Mass Spectrometry Analysis of the sorted MinE07Lib Aptamers

### 5.4.7.1. LC-MS/MS of MinE07Lib Aptamers

Once the method for running the parent aptamer **MinE07** was established, and analysis of the data had been achieved, the top sorted **MinE07Lib** aptamers were analysed. 15 aptamers were chosen at random from the top 170 aptamers (table 5.1).  As can be seen from Table 5.1, only of these hits were above the LOD when analysed b LC-MS/MS and one was borderline. These were MinE07- 41, MinE07-56, MinE07-139 and MinE07-43 respectively.

Table 5.1. Aptamers run on LC-MS/MS and the condition of their data.

| Aptamer | Data Condition |
| --- | --- |
| MinE07-24 | Below LOD |
| MinE07-36 | Below LOD |
| MinE07-41 | Above LOD - data to be analysed |
| MinE07-42 | Below LOD |
| MinE07-43 | MS/MS fragment pattern on edge of LOD – not conclusive data. |
| MinE07-54 | Below LOD |
| MinE07-55 | Below LOD |
| MinE07-56 | Above LOD - data to be analysed |
| MinE07-98 | Below LOD |
| MinE07-111 | Below LOD |
| MinE07-125 | Below LOD |
| MinE07-137 | Below LOD |
| MinE07-139 | Above LOD - data to be analysed |

| MinE07-141 | Below LOD |
|---|---|
| MinE07-153 | Below LOD |

### 5.4.7.2. LC-MS/MS Analysis of MinE07-41

MinE07-41: 5'-fCrGrGrAfU<span style="color:magenta">fU-Ph</span><span style="color:magenta">fU-Ph</span>rArAfUfCrGfCfCrG<span style="color:blue">fU-I</span>rArGrArArArArGfCrAfUrGfUfCrArAr

ArGfCfCrGrGrArAfCfCrG<span style="color:magenta">fU-Ph</span>fCfC-PCL-3'
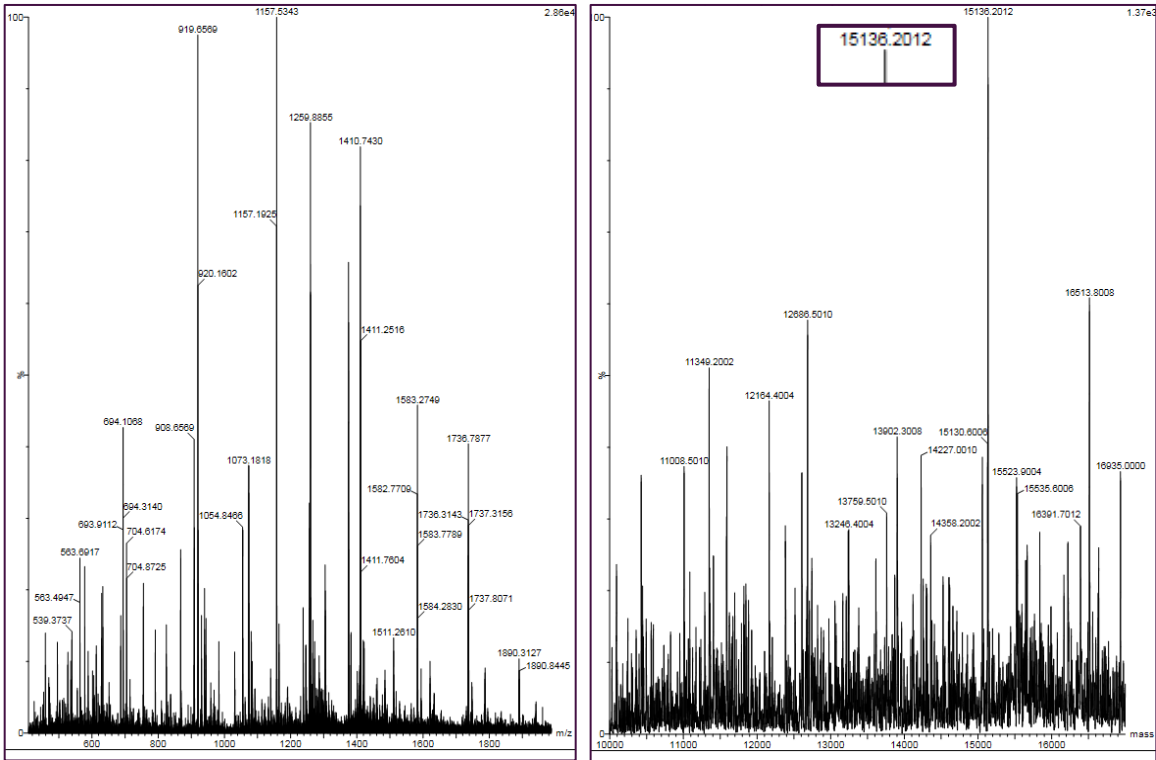


Figure 5.60. LC-MS spectra of MinE07-41 and predicted mass.

Figure 5.60 shows the mass spectrum of **MinE07-41** at RT of 36.108 minutes. The mass is predicted to be 15,136.2 Da which suggests that some nucleotides have been capped off as the MinE07Lib mass range is 15677-16693.2 Da. Upon analysis of the MS/MS fragment data it shows that during the synthesis the aptamer had been capped at base 45 (fC) meaning that the 5'-rGrGrA sequence was missing from the end as there was no fragment pattern for this section. The fragment data shows that fU2/3/8 are <span style="color:magenta">fU-Ph</span> and fU5 is <span style="color:blue">fU-I.</span> Not all of the fragments have

data for every nucleoside that it is expected. This is due to the sample being on the edge of the LOD and some of the peaks were below the threshold and so not visible. As there are data points missing, this being the correct structure will be validated by the protein binding assays. Figure 5.61 shows all the discovered fragments in **MinE07-41.** This figure displays all the overlapping fragments, 3rd fU-Ph, 4th fU, 5th fU-I and 8th fU-Ph can all be seen in more than one fragment, making this uridine identifications more reliable, however not all uridines are displayed in more than one fragment. All individual fragment data are shown below in figures 5.62-5.75.

5'-fCrGrGrA | **fUfU-PhfU-PhrArA** | **fUfCrGfCfC** | rGfU-IrArGrArArArArGfCrA | **fUrGfUfCrA** | rArArGfCfCrG | **rGrArAfCfCrGfU-PhfC** | fC-PCL-3'

| **fU-PhrArAfUfCrG** |                                               | **fCfCrGfU-PhfC** |
                    | **rGfCfCrGfU-IrA** |
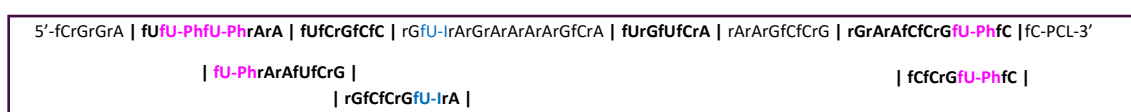
Figure 5.61. MinE07-41 Fragment Pattern. Fragments are shown by the bold sections.

### 5.4.7.2.1. Fragment 1 of MinE07-41

5'-fCrGrGrA | **fUfU-PhfU-PhrArA |** fUfCrGfCfCrGfU-IrArGrArArArArGfCrAfUrGfUfCrArArArGfCf

CrGrGrArAfCfCrGfU-PhfCfC-PCL-3'
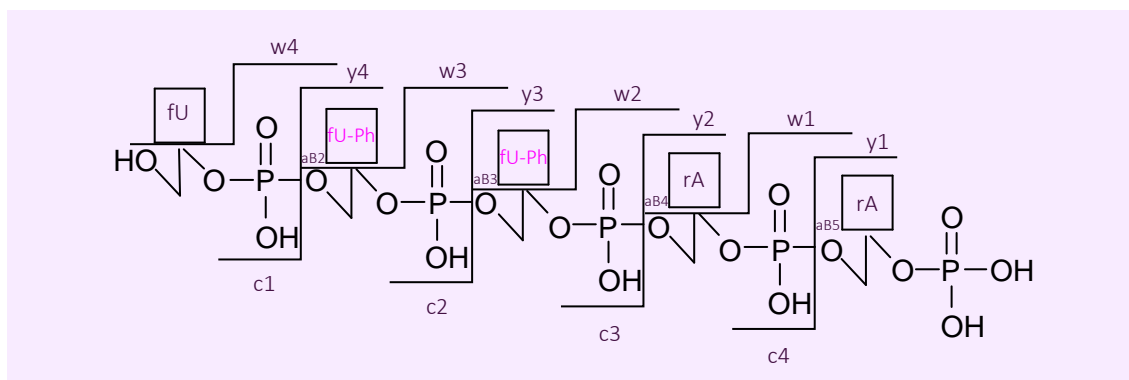


Figure 5.62. Fragment 1 of MinE07-41 5'-3'.
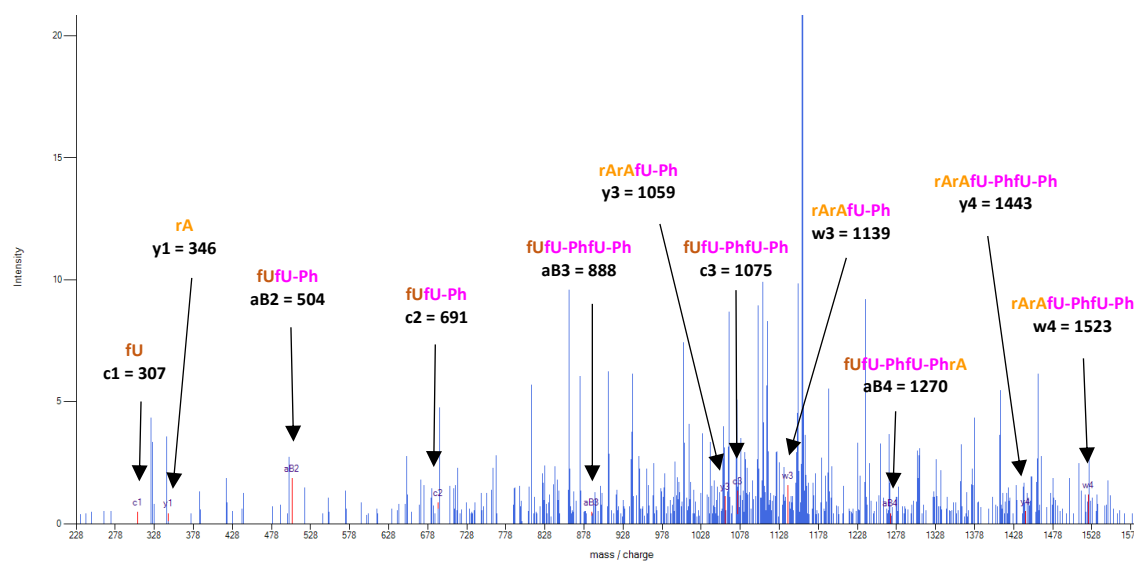


Figure 5.63. LC-MS/MS Fragment 1 Data of MinE07-41.

### 5.4.7.2.2. Fragment 2 of MinE07-41

5'-fCrGrGrAfUfU-Ph | **fU-PhrArAfUfCrG** | fCfCrGfU-IrArGrArArArArGfCrAfUrGfUfCrArArArGfCf

CrGrGrArAfCfCrGfU-PhfCfC-PCL-3'
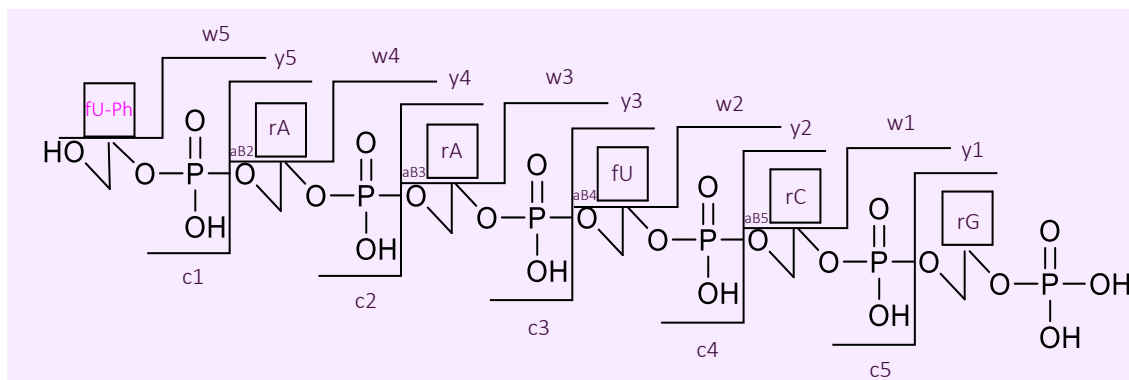


Figure 5.64. Fragment 2 of MinE07-41 5'-3'.
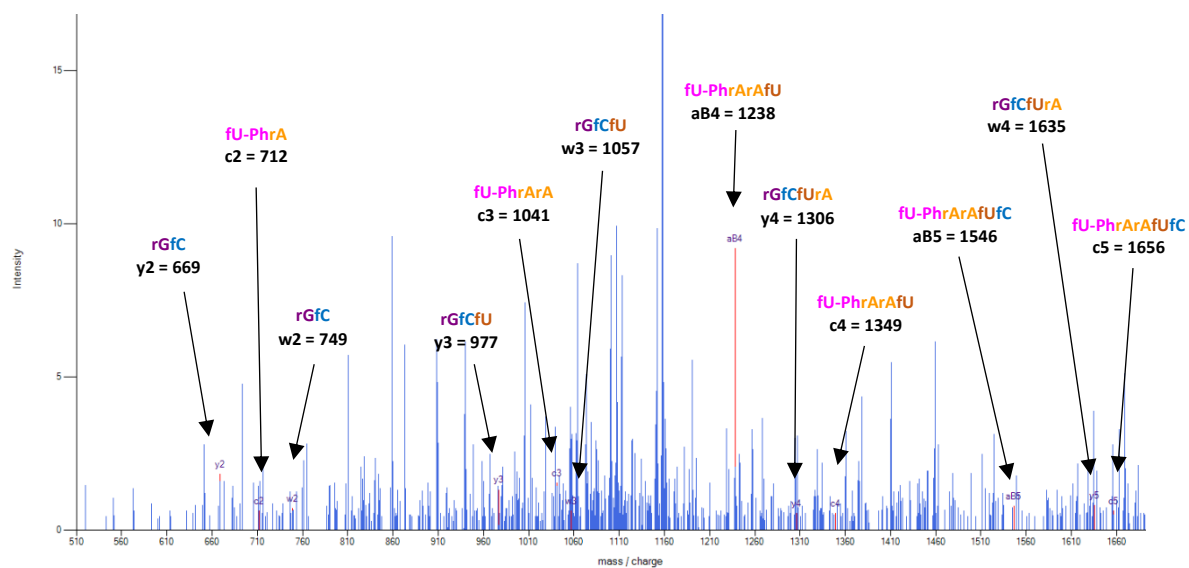


Figure 5.65. LC-MS/MS Fragment 2 Data of MinE07-41.

### 5.4.7.2.3. Fragment 3 of MinE07-41

5'-fCrGrGrAfU**fU-Ph**fU-PhrArA **| fUfCrGfCfC |** rG**fU-I**rArGrArArArArGfCrAfUrGfUfCrArArArGfCf

CrGrGrArAfCfCrG**fU-Ph**fCfC-PCL-3'
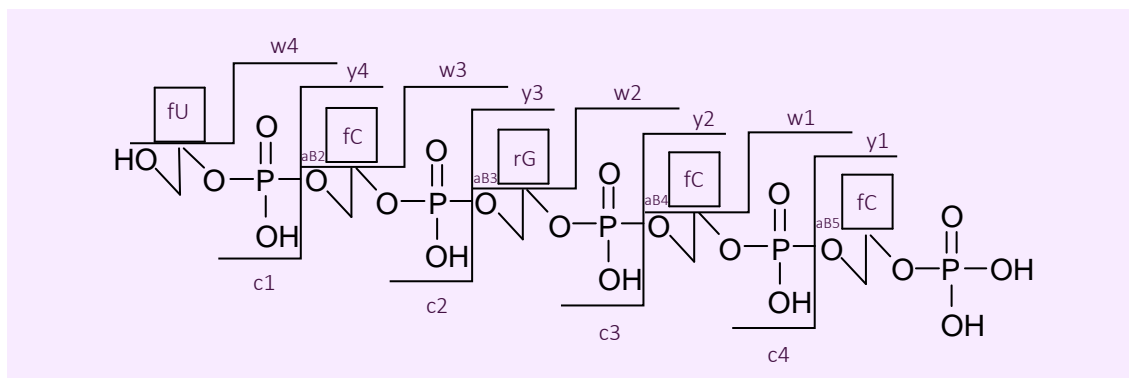


Figure 5.66. Fragment 3 of MinE07-41 5'-3'.
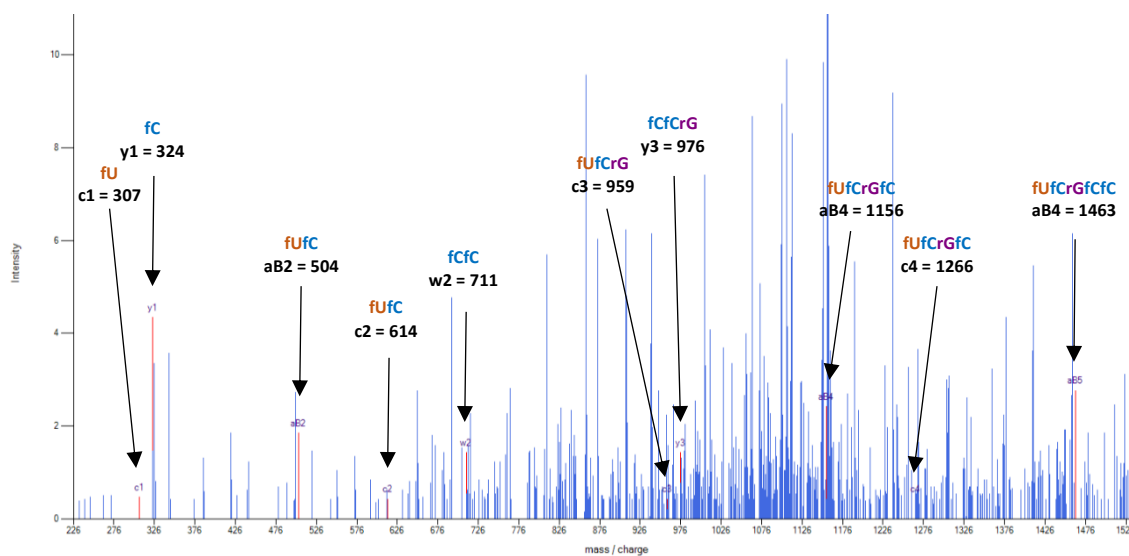


Figure 5.67. LC-MS/MS Fragment 3 Data of MinE07-41.

### 5.4.7.2.4. Fragment 4 of MinE07-41

5'-fCrGrGrAfU<span style="color:magenta">fU-PhfU-Ph</span>rArAfUfC **| rGfCfCrG<span style="color:blue">fU-Ir</span>A |** rGrArArArArGfCrAfUrGfUfCrArArArGfCf

CrGrGrArAfCfCrG<span style="color:magenta">fU-Ph</span>fCfC-PCL-3'
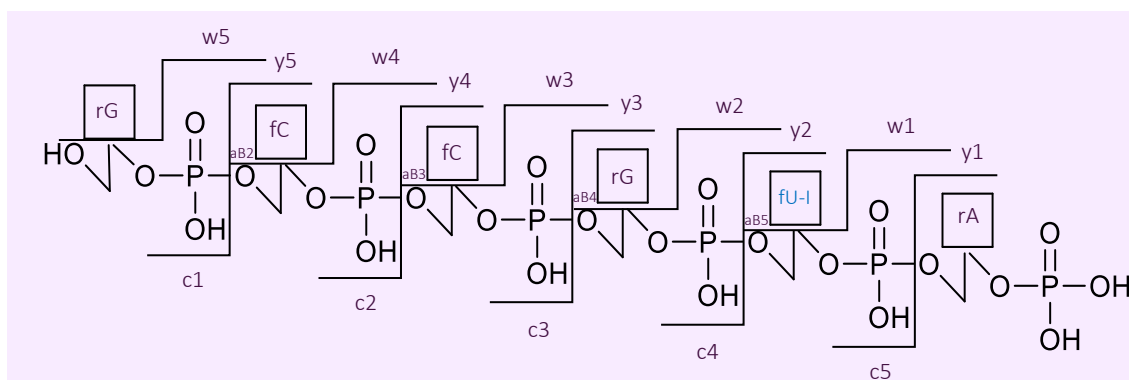


Figure 5.68. Fragment 4 of MinE07-41 5'-3'.
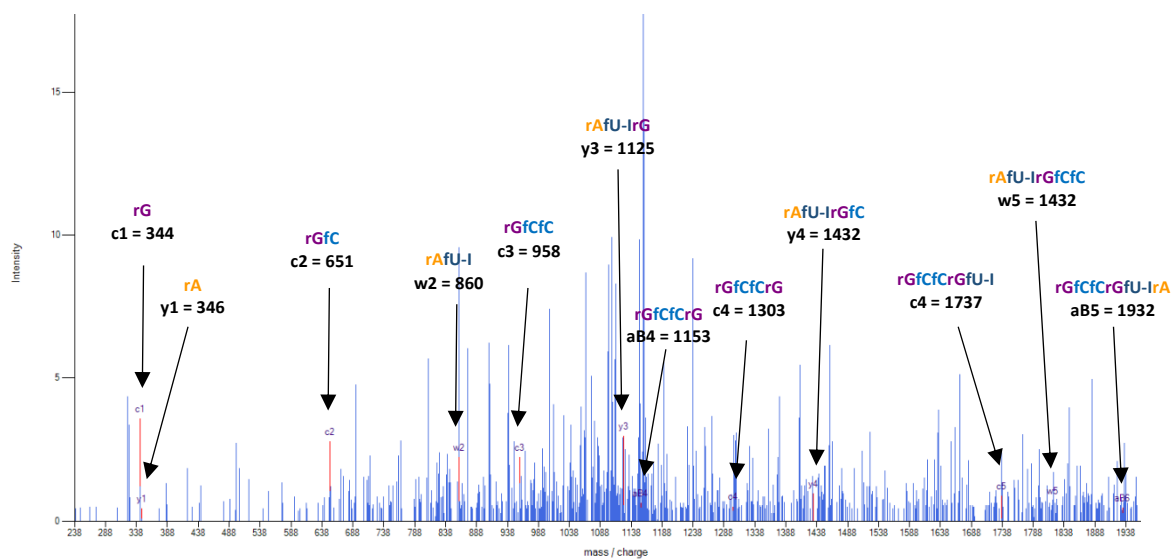


Figure 5.69. LC-MS/MS Fragment 4 Data of MinE07-41.

5'-fCrGrGrAfUfU-PhfU-PhrArAfUfCrGfCfCrGfU-IrArGrArArArArGfCrA **| fUrGfUfCrA |** rArArGfCf

CrGrGrArAfCfCrGfU-PhfCfC-PCL-3'
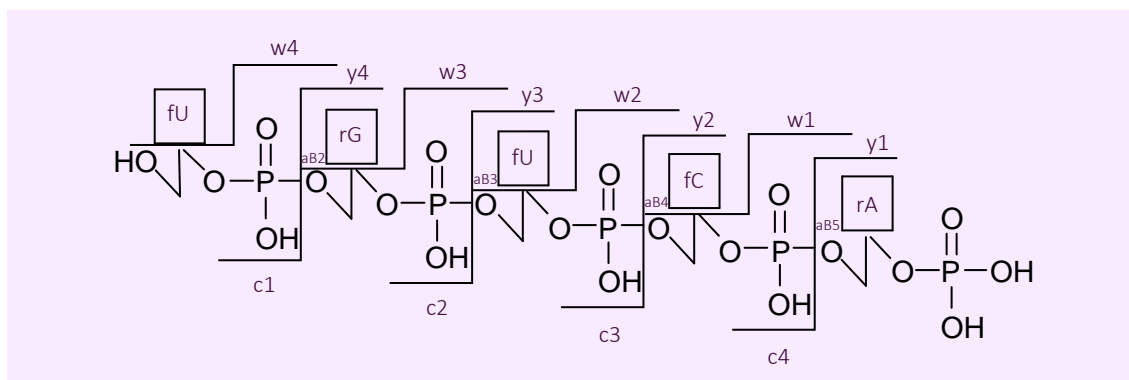


Figure 5.70. Fragment 5 of MinE07-41 5'-3'.
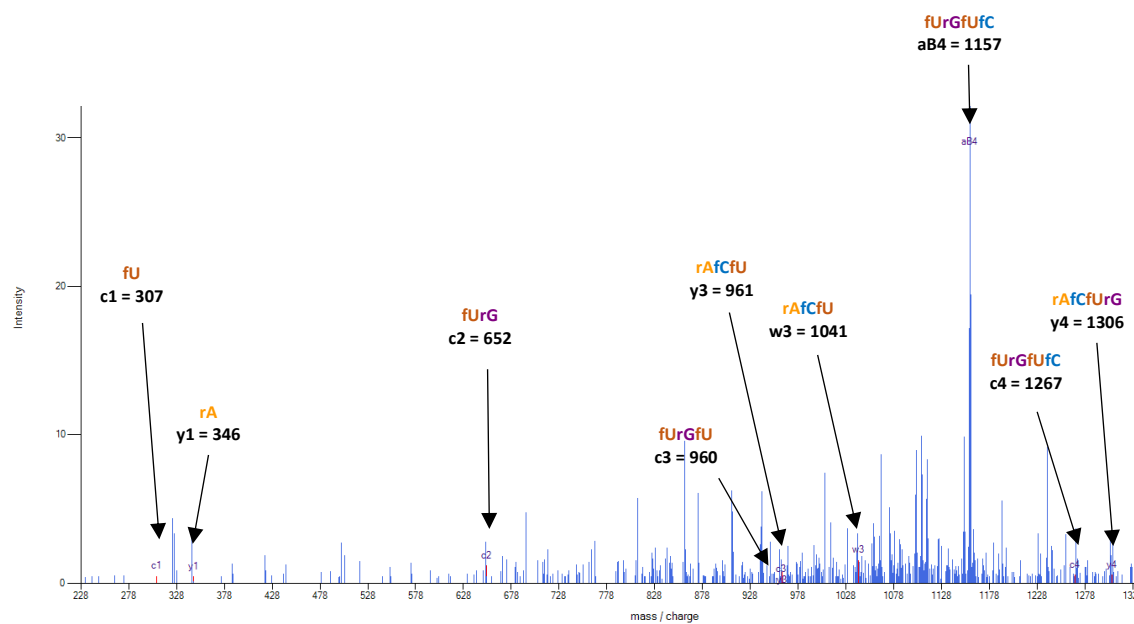


Figure 5.71. LC-MS/MS Fragment 5 Data of MinE07-41.

5'-fCrGrGrAfUfU-PhfU-PhrArAfUfCrGfCfCrGfU-IrArGrArArArArGfCrAfUrGfUfCrArArArGfCfCrG **|**

**rGrArAfCfCrGfU-PhfC |** fC-PCL-3'
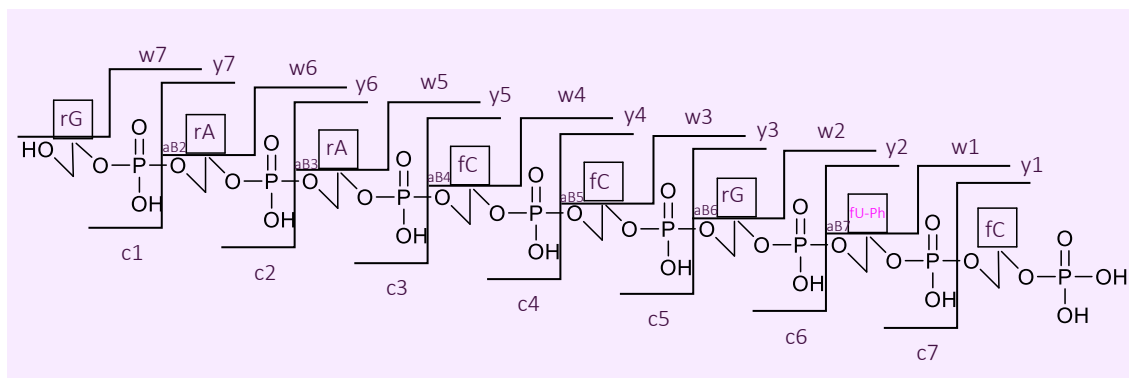


Figure 5.72. Fragment 6 of MinE07-41 5'-3'.
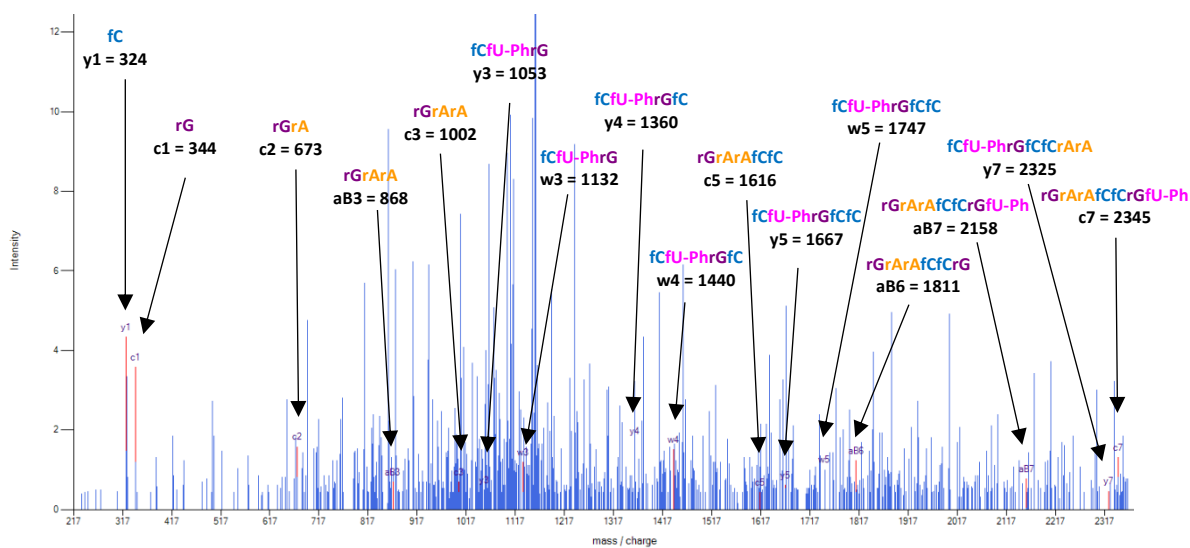


Figure 5.73. LC-MS/MS Fragment 6 Data of MinE07-41.

### 5.4.7.2.7. Fragment 7 of MinE07-41

5'-fCrGrGrAfUfU-PhfU-PhrArAfUfCrGfCfCrGfU-IrArGrArArArArGfCrAfUrGfUfCrArArArGfCfCrGr

GrArA | **fCfCrGfU-PhfC** | fC-PCL-3'
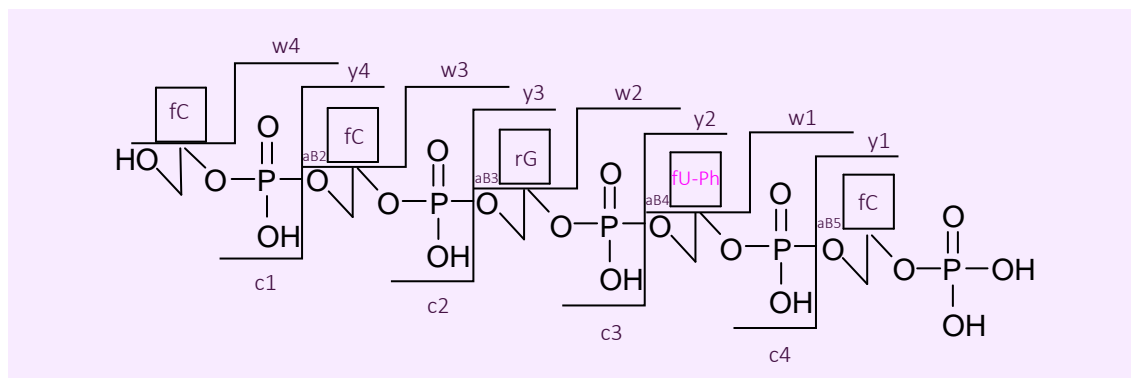


Figure 5.74. Fragment 7 of MinE07-41 5'-3'.
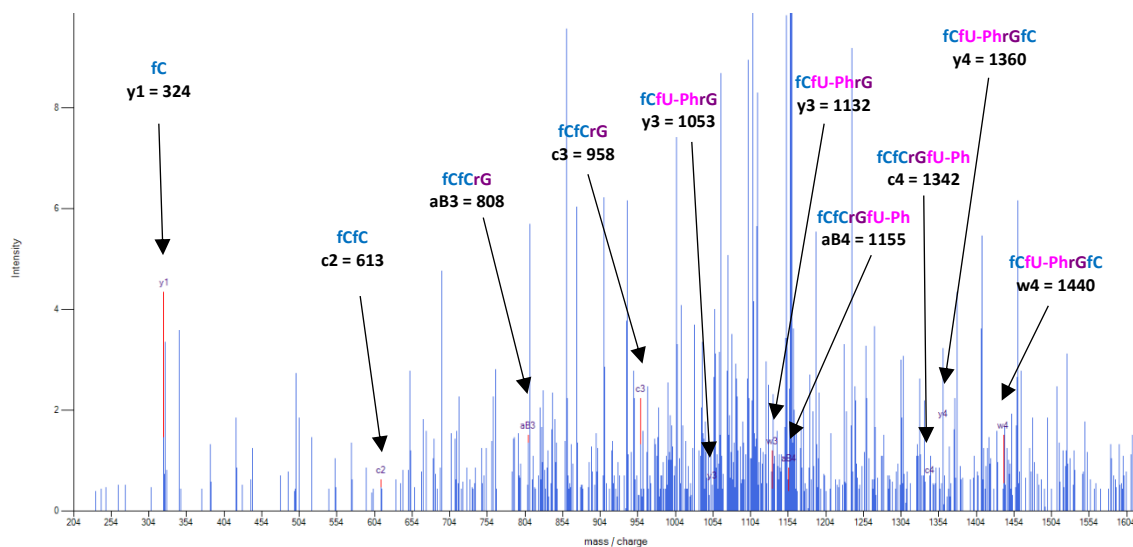


Figure 5.75. LC-MS/MS Fragment 7 Data of MinE07-41.

The data of the fragment patterns of **MinE07-41** are not as complete as the parent MinE07. This

is down to the low concentration of the sample. The accuracy of the sequence will be confirmed

by the protein binding affinity assays (chapter 6).

### 5.4.7.3. LC-MS/MS Analysis of MinE07-56

MinE07-56: 5'-rGrGrAfCrGrGrAfUfU-PhfU-PhrArAfU-IfCrGfCfCrGfU-PhrArGrArArArArGfCrAfUr

GfUfCrArArArGfCfCrGrGrArAfCfCrGfUfCfC-PCL-3'

Figure 5.76 shows the mass spectrum of **MinE07-56** at RT of 38.954 minutes. The mass is

predicted to be 16,150.9 Da which is inside the MinE07Lib mass range 15677-16693.2 Da. Upon

analysis of the MS/MS fragment data it shows that **MinE07-56** is full length. The fragment data

shows that fU2/3/5 are fU-Ph and fU4 is fU-I. Most of the fragments have a lot of data points

missing because of these peaks being below the threshold because the concentration of the

sample is close to the LOD. As there are data points missing, this being the correct structure will

be validated by the protein binding assays. Fragment pattern analysis can be found below.
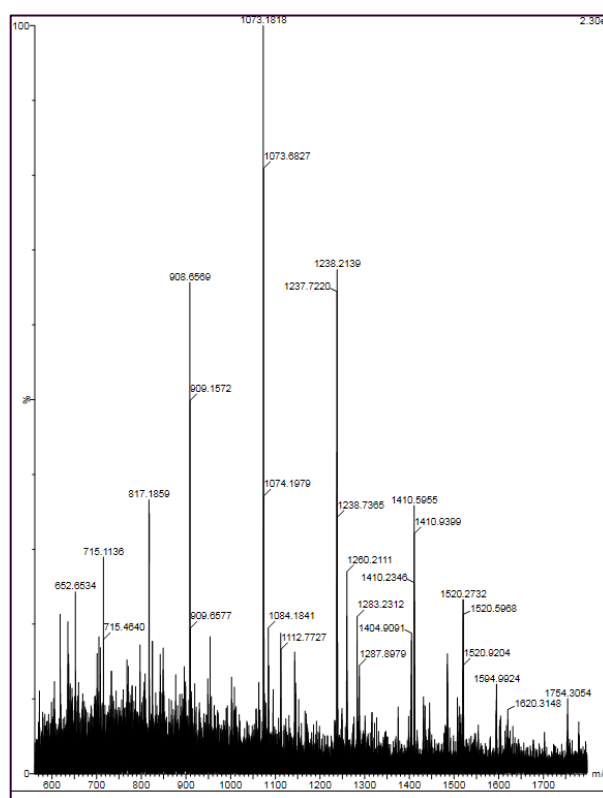


Figure 5.76. LC-MS spectra of MinE07-56.

Figure 5.77 shows all the discovered fragments in **MinE07-56.** This figure displays all the

overlapping fragments, 1st fU, 2nd fU-Ph, 3rd fU-Ph, 5th fU-Ph, 6th fU, 7th fU-I and 8th fU can all be

seen in more than one fragment, making this uridine identifications more reliable, however not all uridines are displayed in more than one fragment. All individual fragment data are shown below in figures 5.78-5.87.

5'-rGrGrAfCrGrG **| rAfUfU-PhfU-PhrA |** rAfU-IfCrGfCfCrGfU-PhrA **| rGrArArArA |** rG **| fCrAfUrGfU-IfC |** rArArArGfCfCrGrG **| rArAfCfCrGfU |** fCfC-PCL-3'
**| fUfU-PhfU-PhrArA |** **| fCrGfU-PhrArGrA |** **| rAfUrGfU-IfC |** **| rAfCfCrGfUfC |**
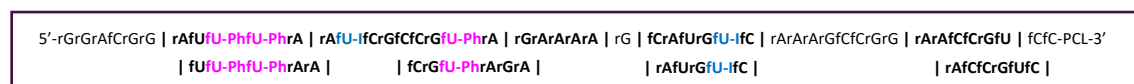
Figure 5.77. MinE07-56 Fragment Pattern. Fragments are shown by the bold sections.

### 5.4.7.3.1. Fragment 1 of MinE07-56

5'-rGrGrAfCrGrG **| rAfUfU-PhfU-PhrA |** rAfU-IfCrGfCfCrGfU-PhrArGrArArArArArGfCrAfUrGfU-IfCr

ArArArGfCfCrGrGrArAfCfCrGfUfCfC-PCL-3'
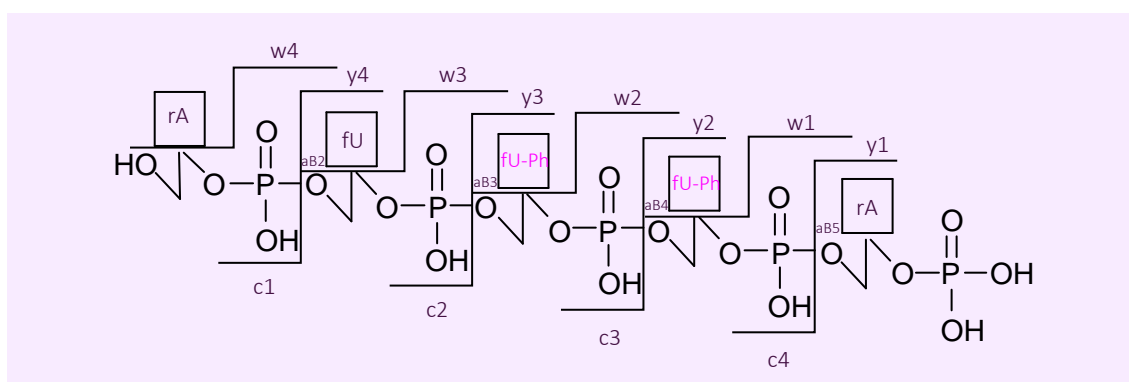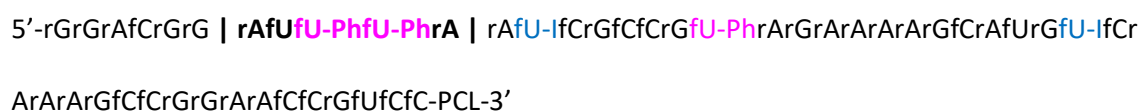


Figure 5.70. Fragment 1 of MinE07-56 5'-3'.
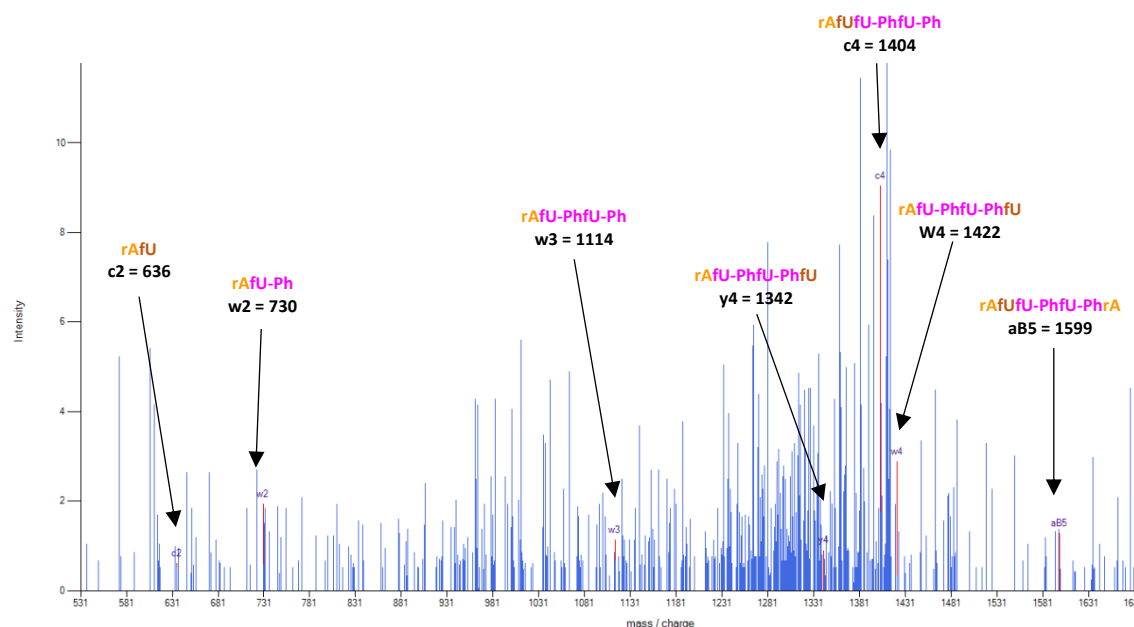


Figure 5.71. LC-MS/MS Fragment 1 Data of MinE07-56.

313

5'-rGrGrAfCrGrGrA | **fUfU-PhfU-PhrArA** | fU-IfCrGfCfCrGfU-PhrArGrArArArArGfCrAfUrGfU-IfCr

ArArArGfCfCrGrGrArAfCfCrGfUfCfC-PCL-3'
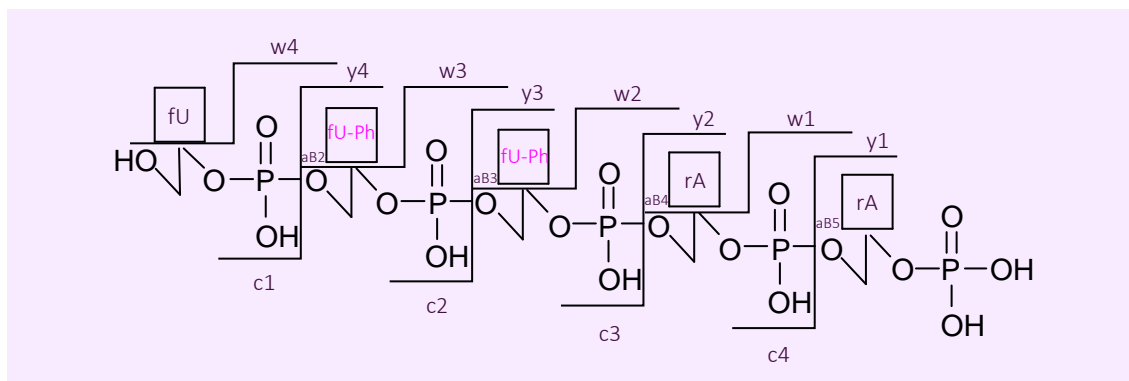


Figure 5.72. Fragment 2 of MinE07-56 5'-3'.
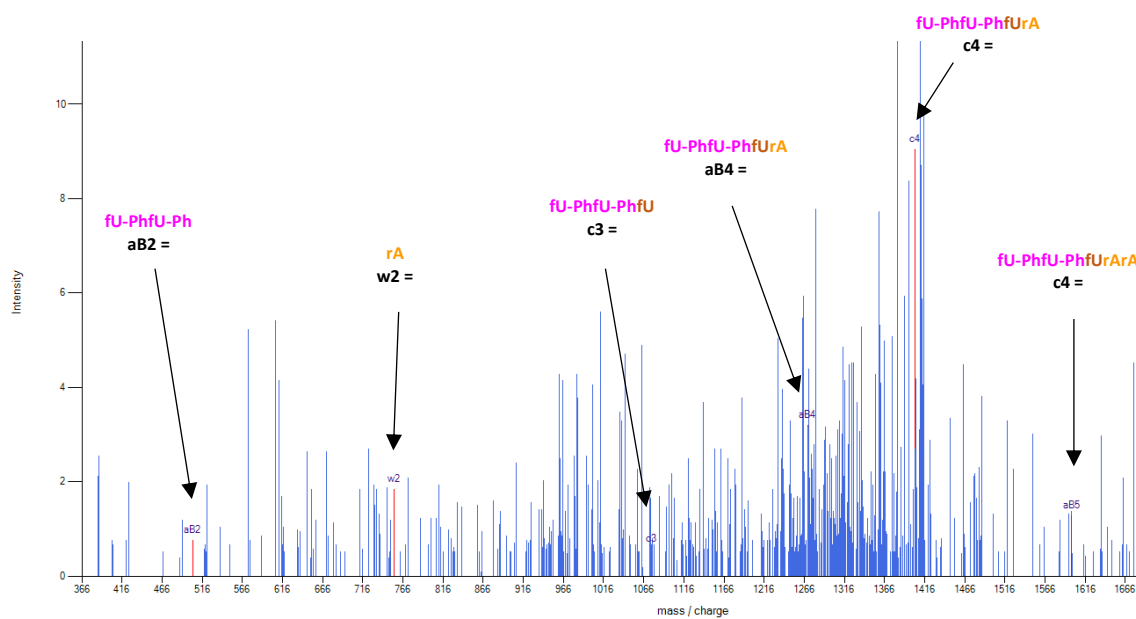


Figure 5.73. LC-MS/MS Fragment 2 Data of MinE07-56.

### 5.4.7.3.3. Fragment 3 of MinE07-56

5'-rGrGrAfCrGrGrAfUfU-PhfU-PhrA **| rAfU-IfCrGfCfCrGfU-Ph rA |** rGrArArArArGfCrAfUrGfU-IfCr
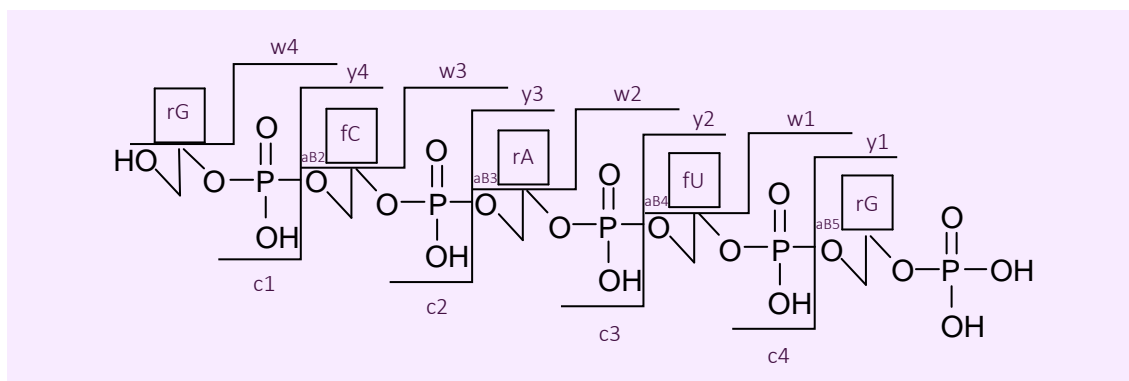
ArArArGfCfCrGrGrArAfCfCrGfUfCfC-PCL-3'
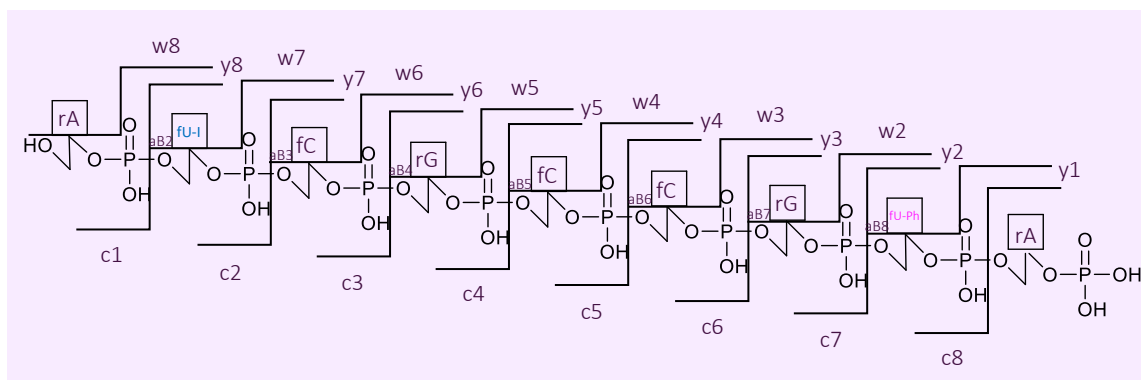


Figure 5.74. Fragment 3 of MinE07-56 5'-3'.
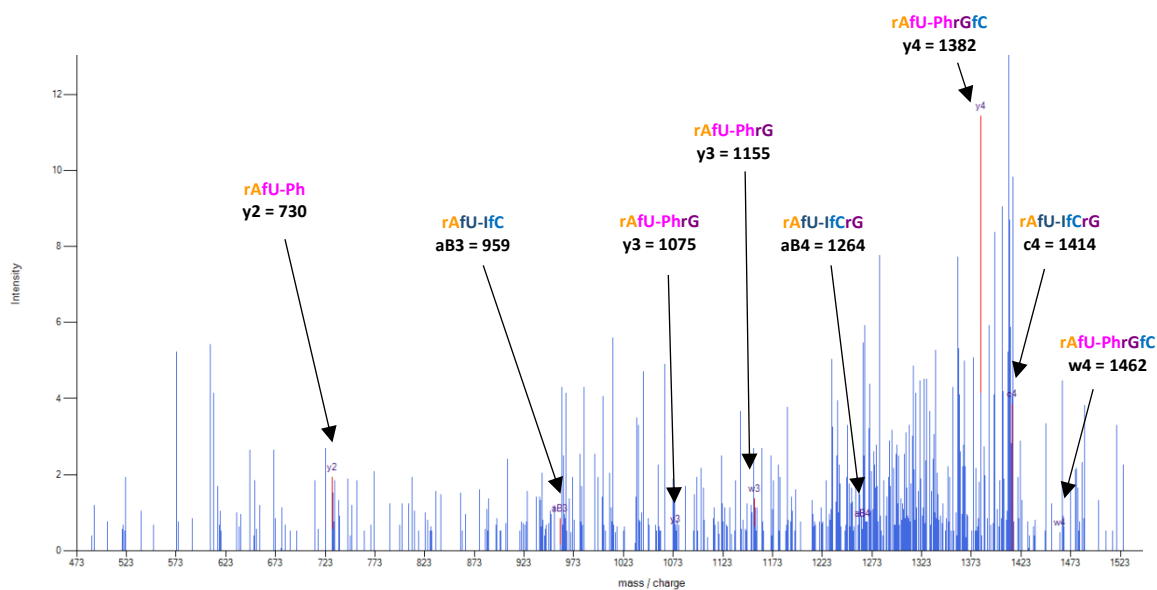


Figure 5.75. LC-MS/MS Fragment 3 Data of MinE07-56.

### 5.4.7.3.4. Fragment 4 of MinE07-56

5'-rGrGrAfCrGrGrAfUfU-PhfU-PhrArAfU-IfCrGfC **| fCrGfU-PhrArGrA |** rArArArGfCrAfUrGfU-IfCr

ArArArGfCfCrGrGrArAfCfCrGfUfCfC-PCL-3'



Figure 5.76. Fragment 4 of MinE07-56 5'-3'.
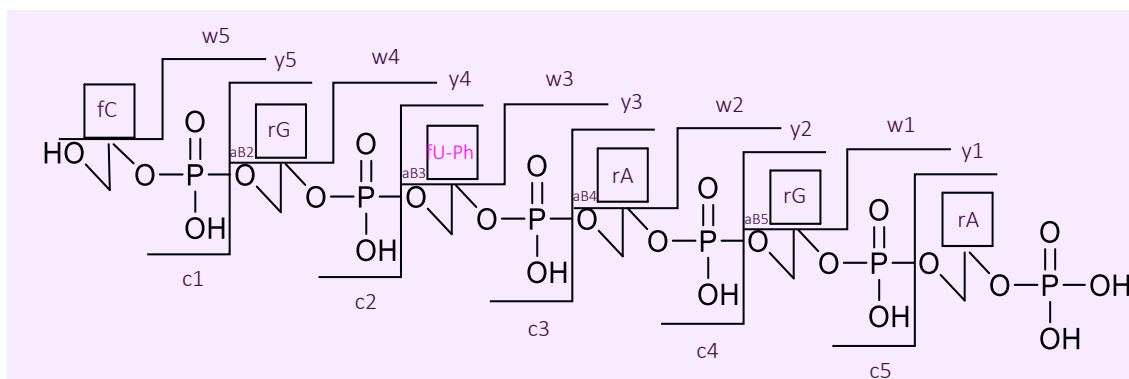


Figure 5.77. LC-MS/MS Fragment 4 Data of MinE07-56.

### 5.4.7.3.5. Fragment 5 of MinE07-56

5'-rGrGrAfCrGrGrAfUfU-PhfU-PhrArAfU-IfCrGfCfCrGfU-PhrA **| rGrArArArA |** rGfCrAfUrGfU-IfCr

ArArArGfCfCrGrGrArAfCfCrGfUfCfC-PCL-3'
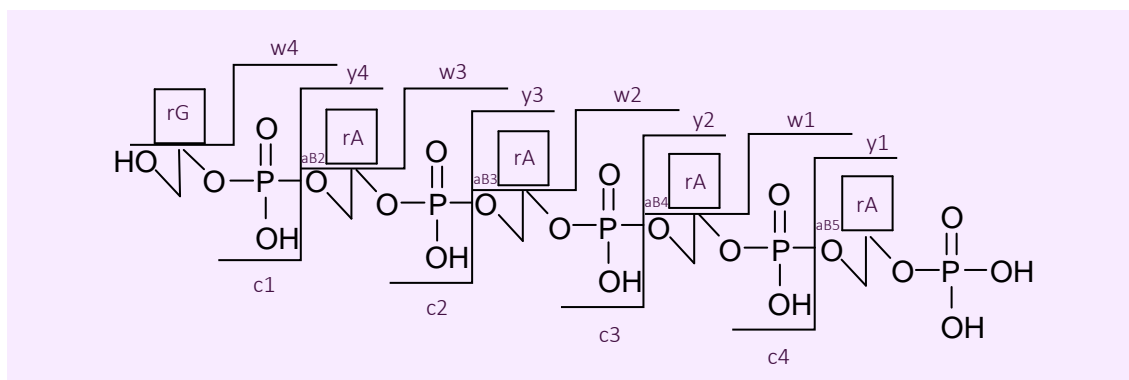


Figure 5.78. Fragment 5 of MinE07-56 5'-3'.
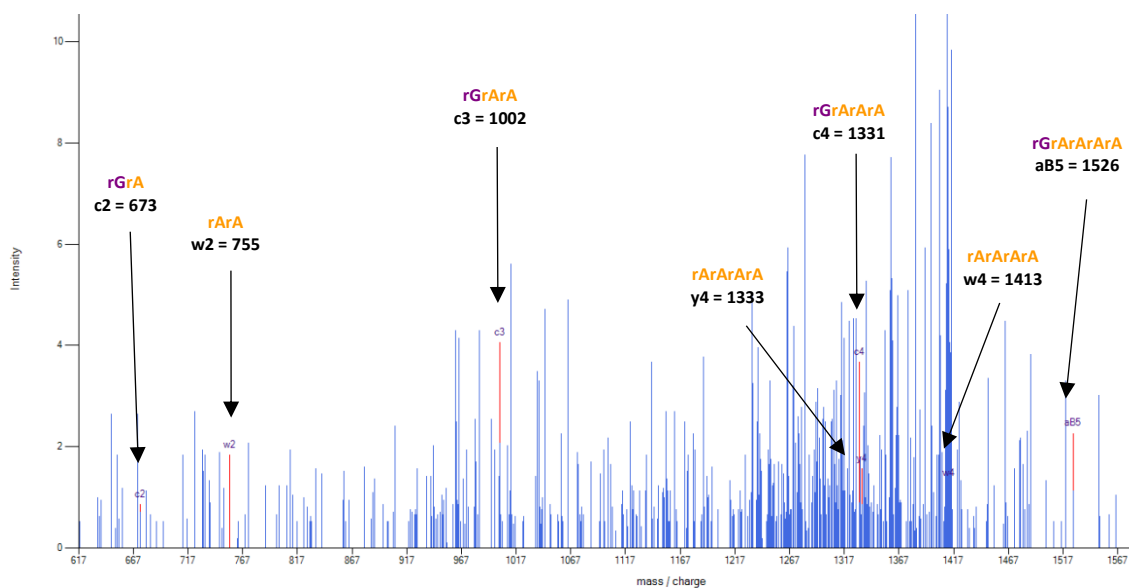


Figure 5.79. LC-MS/MS Fragment 5 Data of MinE07-56.

5'-rGrGrAfCrGrGrAfU<span style="color:magenta">fU-PhfU-Ph</span>rArA<span style="color:blue">fU-I</span>fCrGfCfCrG<span style="color:magenta">fU-Ph</span>rArGrArArArArG **| fCrAfUrGfU-IfC |**

rArArArGfCfCrGrGrArAfCfCrGfUfCfC-PCL-3'
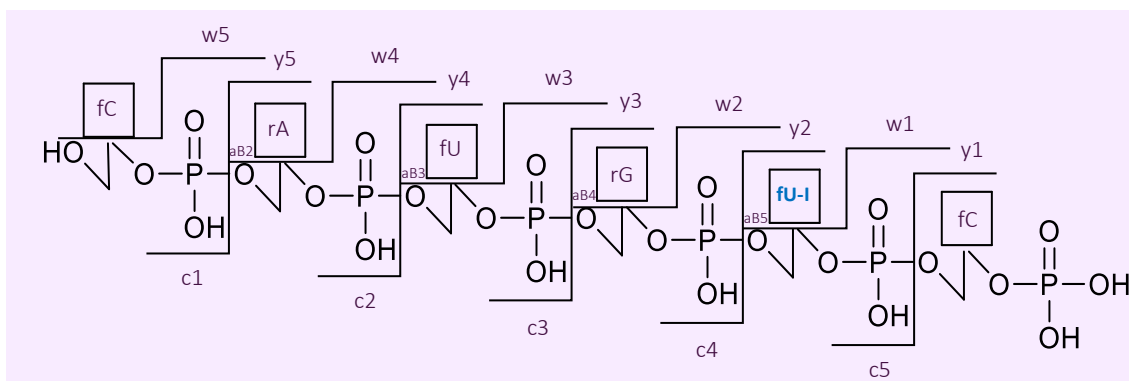


Figure 5.80. Fragment 6 of MinE07-56 5'-3'.
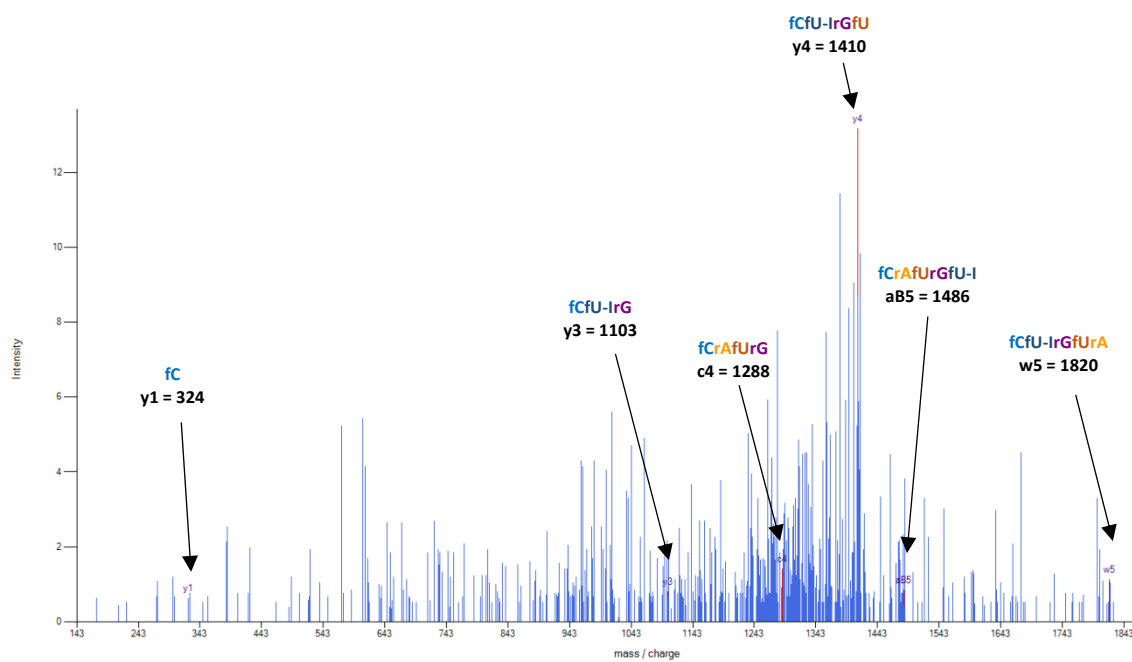


Figure 5.81. LC-MS/MS Fragment 6 Data of MinE07-56.

### 5.4.7.3.7. Fragment 7 of MinE07-56

5'-rGrGrAfCrGrGrAfU<span style="color:magenta">fU-P</span>h<span style="color:magenta">fU-P</span>hrArA<span style="color:blue">fU-I</span>fCrGfCfCrG<span style="color:magenta">fU-P</span>hrArGrArArArArGfC **| rAfUrG<span style="color:blue">fU-I</span>fC |** r

ArArArGfCfCrGrGrArAfCfCrGfUfCfC-PCL-3'
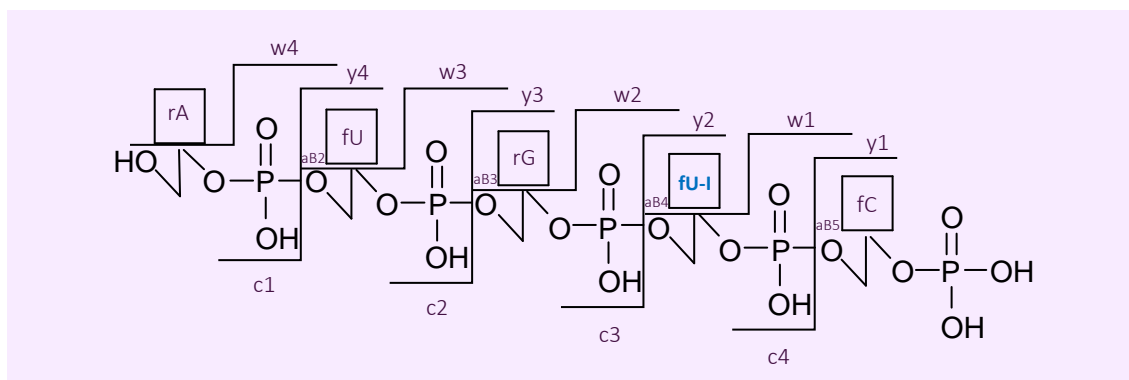


Figure 5.82. Fragment 7 of MinE07-56 5'-3'.
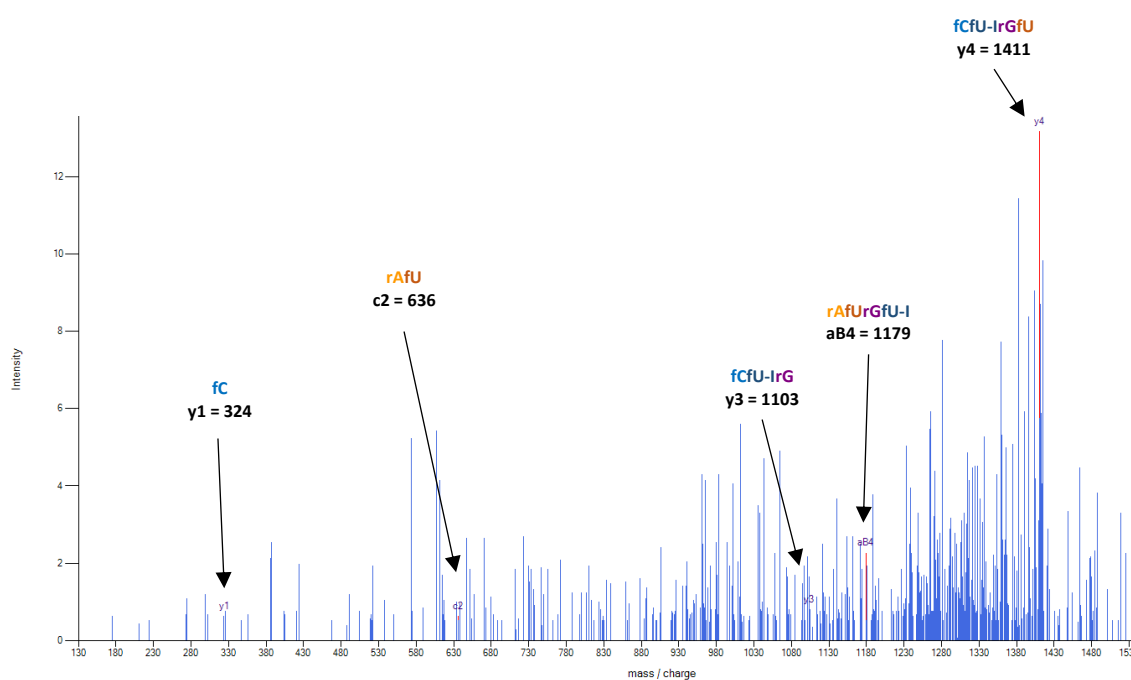


Figure 5.83. LC-MS/MS Fragment 7 Data of MinE07-56.

### 5.4.7.3.8. Fragment 8 of MinE07-56

5'-rGrGrAfCrGrGrAfUfU-PhfU-PhrArAfU-IfCrGfCfCrGfU-PhrArGrArArArArGfCrAfUrGfU-IfCrArAr

ArGfCfCrGrG **| rArAfCfCrGfU |** fCfC-PCL-3'
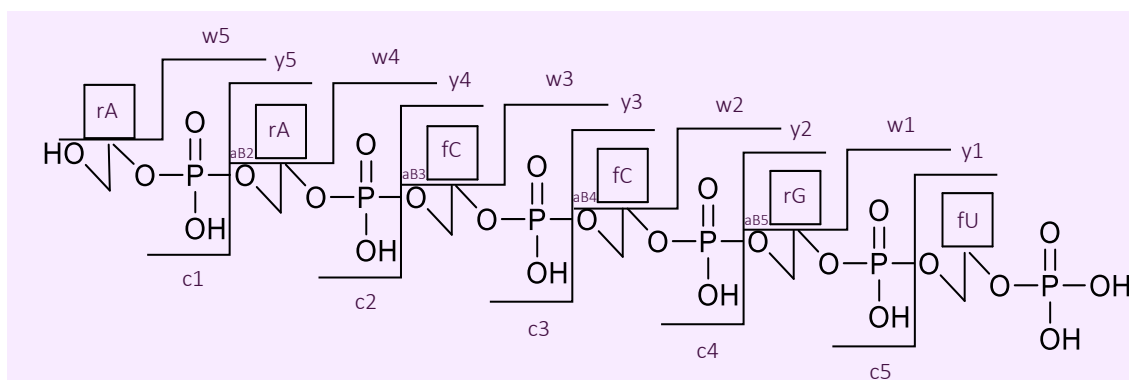


Figure 5.84. Fragment 8 of MinE07-56 5'-3'.
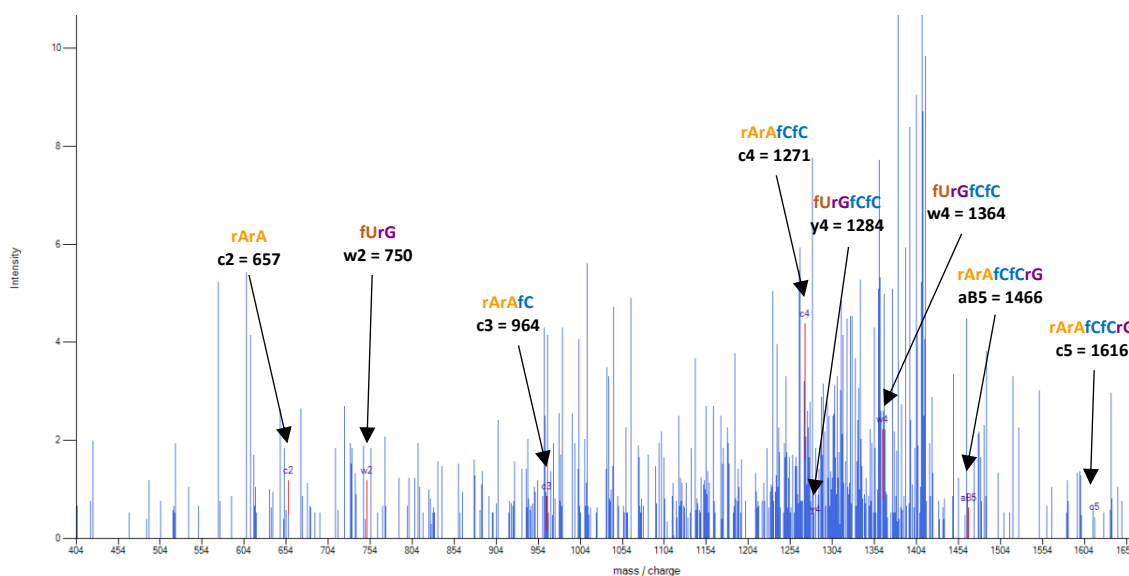


Figure 5.85. LC-MS/MS Fragment 8 Data of MinE07-56.

5'-rGrGrAfCrGrGrAfU<span style="color:magenta">fU-Ph</span>fU-PhrArAf<span style="color:blue">U-I</span>fCrGfCfCrG<span style="color:magenta">fU-Ph</span>rArGrArArArArGfCrAfUrG<span style="color:blue">fU-I</span>fCrArAr

ArGfCfCrGrGrA **| rAfCfCrGfUfC |** fC-PCL-3'
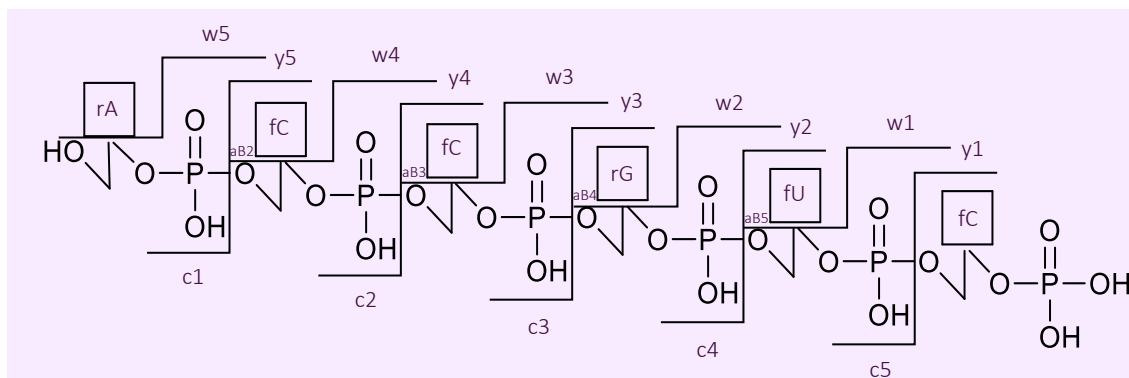


Figure 5.86. Fragment 9 of MinE07-56 5'-3'.
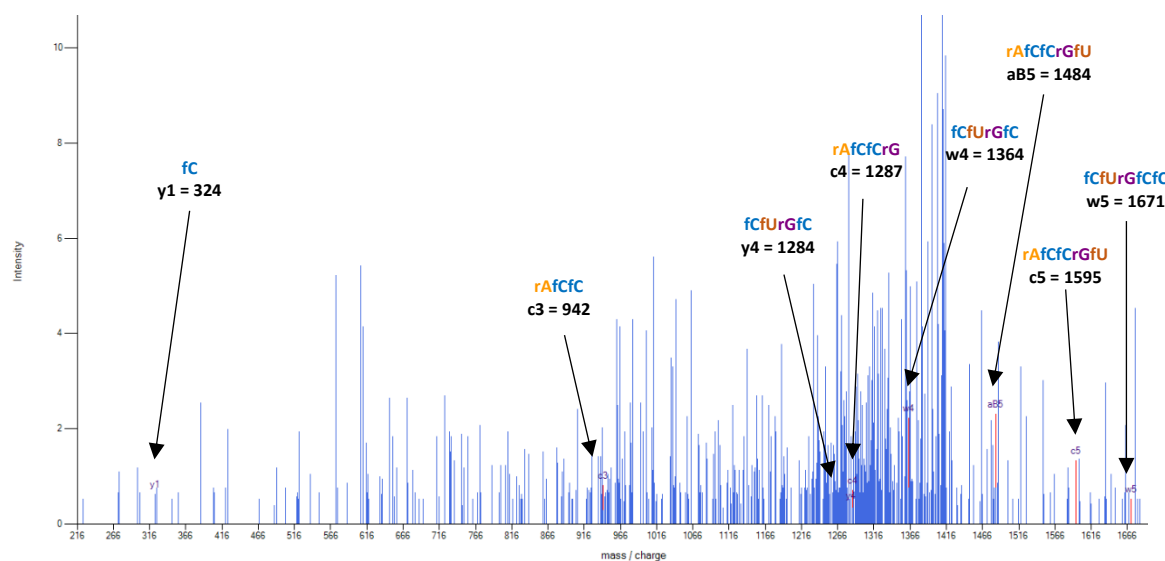


Figure 5.87. LC-MS/MS Fragment 9 Data of MinE07-56.

The data of the fragment patterns of MinE07-56 are not as complete as the parent **MinE07** or **MinE07-41** and **MinE07-139**. The precision of the sequence will be confirmed by the protein binding affinity assays (chapter 6).

### 5.4.7.4. LC-MS/MS Analysis of MinE07-139

MinE07-139: 5'-rGrGrAfU-VifUfUrArAfUfCrGfCfCrGfUrArGrArArArArGfCrAfU-PhrGfU-VifCrArAr

ArGfCfCrGrGrArAfCfCrGfUfCfC-PCL-3'

Figure 5.88 shows the mass spectrum of **MinE07-139** at RT of 40.9 minutes. The mass is

predicted to be 14,536.0 Da which suggests that some nucleotides have been capped off as the

**MinE07Lib** mass range is 15,677-16,693.2 Da. Upon analysis of the MS/MS fragment data it

shows that during the synthesis it had been capped at base 44 (fG) meaning that the 5'-rGrGrAfC

sequencing was missing from the end, as there was no fragment pattern for this section. The

fragment data shows that fU1/7 are fU-Vi and fU6 is fU-Ph.  Most of the fragments have all their

data points, this sample being well within the LOD compared with **MinE07-41** and **MiNE07-56**.

This being the correct structure will be validated by the protein binding assays. Fragment

pattern analysis can be found below.
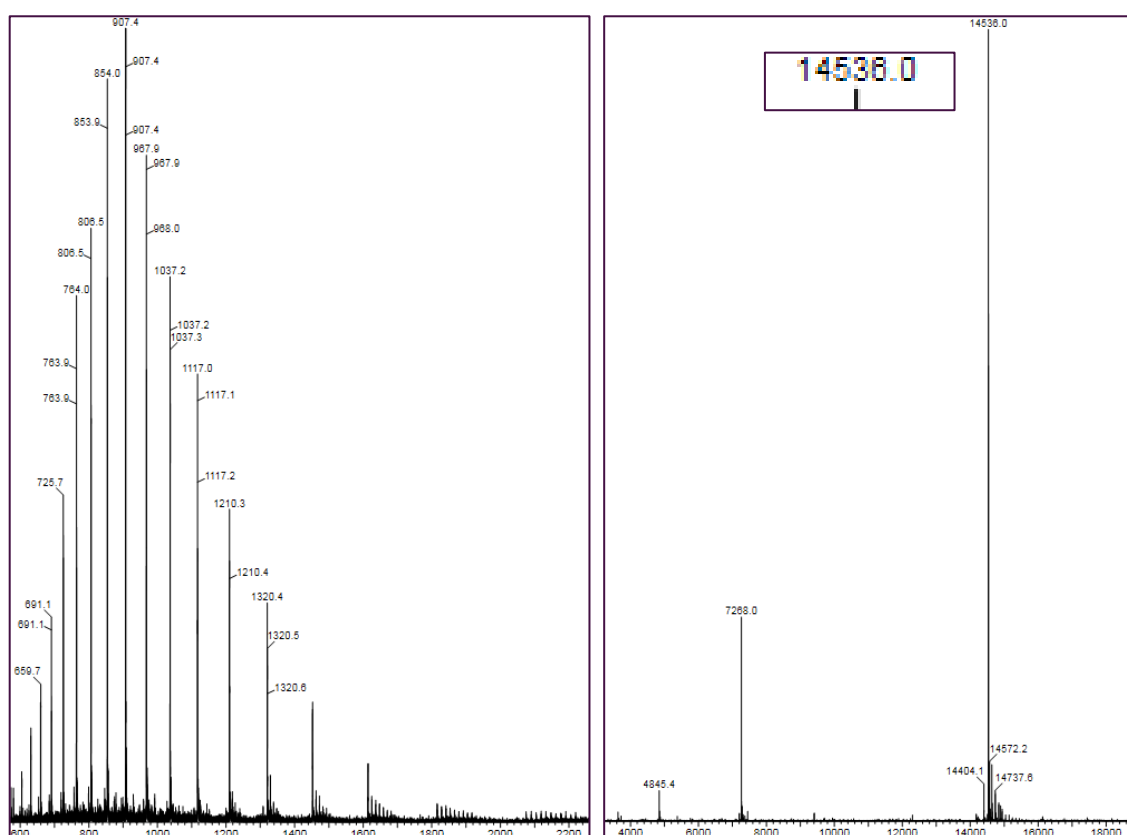


Figure 5.88. LC-MS spectra of MinE07-139 and predicited mass

Figure 5.89 shows all the discovered fragments in **MinE07-139.** This figure displays all the overlapping fragments, 1st fU-Vi, 2nd fU, 3rd fU, 4th fU, 5th fU, can all be seen in more than one fragment, however not all uridines are displayed in more than one fragment. All individual fragment data are shown below in figures 5.90-5.111.

5'-rG | **rGrAfU-VifUfU** | **rArAfUfCrGfC** | fCrGfU | **rArGrArArArA** | rGfC | **rAfU-PhrGfU-VifCrA** | rArArG | **fCfCrGrGrArA** | fC | **fCrGfUfCfC-PCL** | -3
| **rAfU-VifUfUrArA** |            | **rGfCfCrGfUrA** |                             | **fCrArArArG** |              | **rGrArAfCfC** |
                                    | **rAfUfCrGfC** |

Figure 5.89. MinE07-139 Fragment Pattern. Fragments are shown by the bold sections.

### 5.4.7.4.1. Fragment 1 of MinE07-139

5'-rG | **rGrAfU-VifUfU** | rArAfUfCrGfCfCrGfUrArGrArArArArArGfCrAfU-PhrGfU-VifCrArArArGfCfCr GrGrArAfCfCrGfUfCfC-PCL-3'



Figure 5.90. Fragment 1 of MinE07-139 5'-3'.



Figure 5.91. LC-MS/MS Fragment 1 Data of MinE07-139.

## 5.4.7.4.2. Fragment 2 of MinE07-139

5'-rGrG **| rAfU-VifUfUrArA |** fUfCrGfCfCrGfUrArGrArArArArGfCrAfU-PhrGfU-VifCrArArArGfCfC

rGrGrArAfCfCrGfUfCfC-PCL-3'
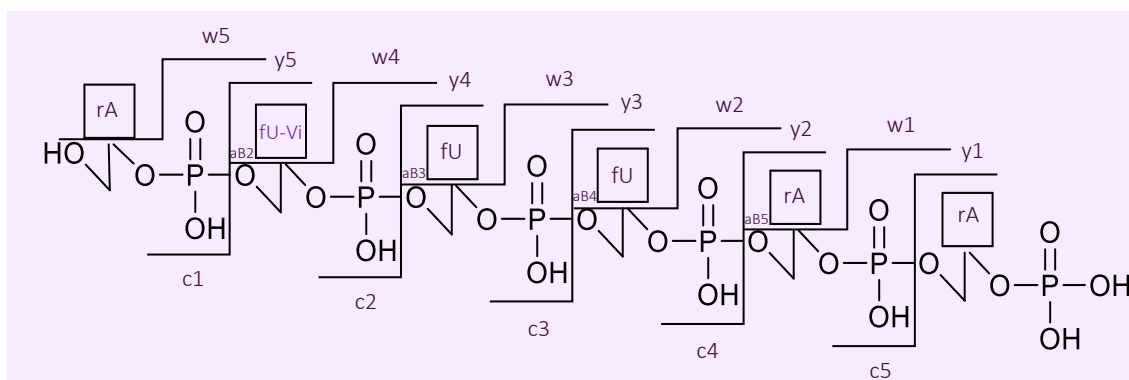


Figure 5.92. Fragment 2 of MinE07-139 5'-3'.
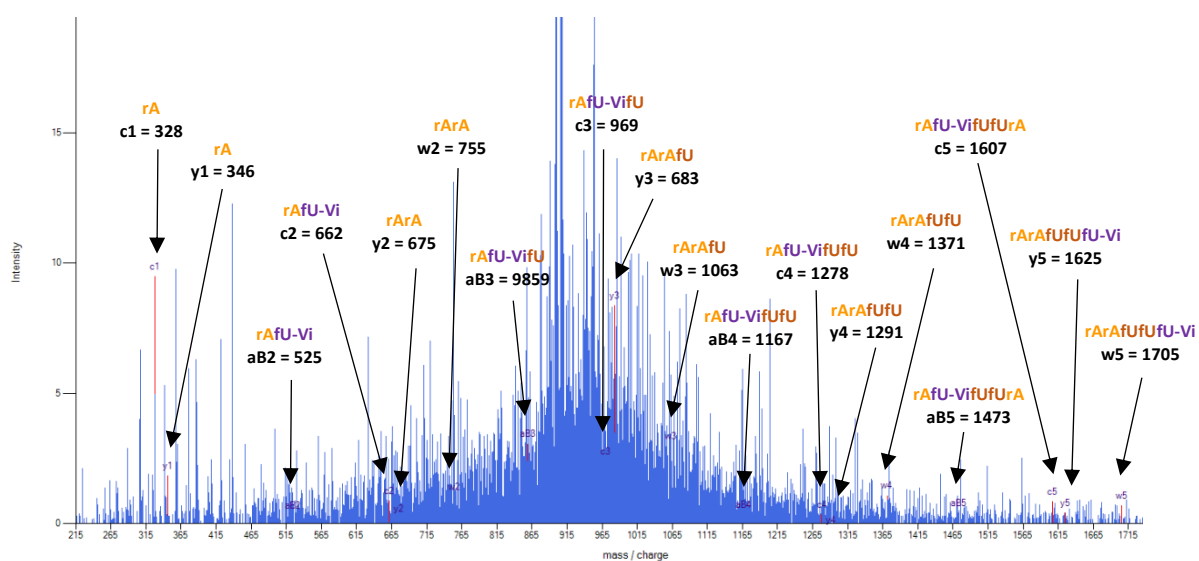


Figure 5.93. LC-MS/MS Fragment 2 Data of MinE07-139.

### 5.4.7.4.3. Fragment 3 of MinE07-139

5'-rGrGrAfU-VifUfU **| rArAfUfCrGfC |** fCrGfUrArGrArArArArGfCrAfU-PhrGfU-VifCrArArArGfCfCr

GrGrArAfCfCrGfUfCfC-PCL-3'
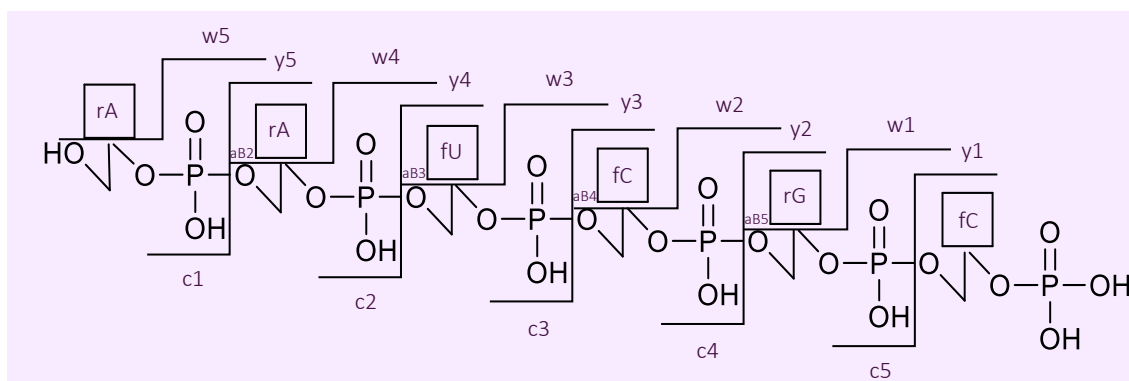


Figure 5.94. Fragment 3 of MinE07-139 5'-3'.
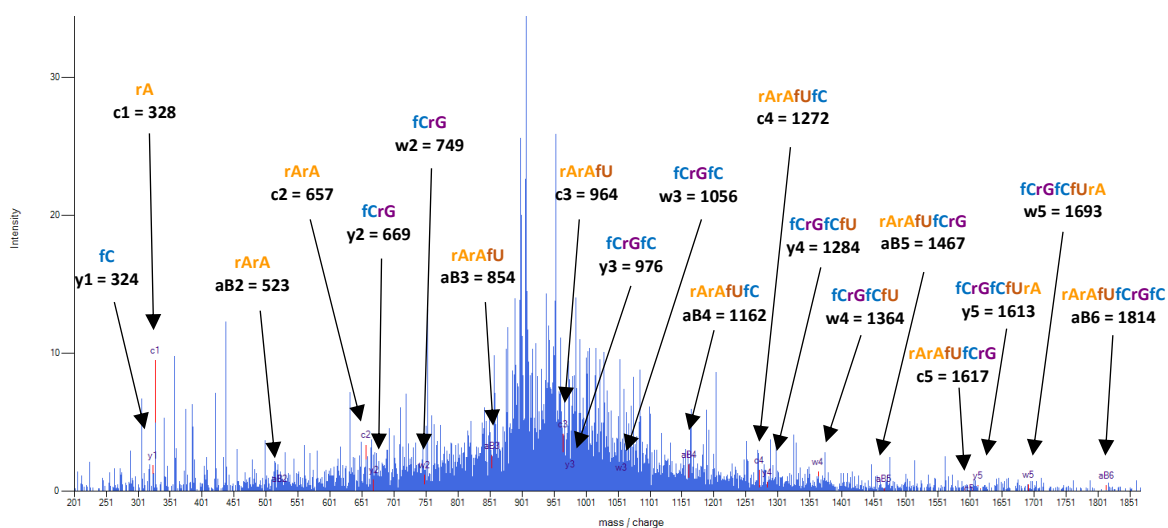


Figure 5.95. LC-MS/MS Fragment 3 Data of MinE07-139.

### 5.4.7.4.4. Fragment 4 of MinE07-139

5'-rGrGrAfU-VifUfUrA **| rAfUfCrGfC |** fCrGfUrArGrArArArArGfCrAfU-PhrGfU-VifCrArArArGfCfCr

GrGrArAfCfCrGfUfCfC-PCL-3'
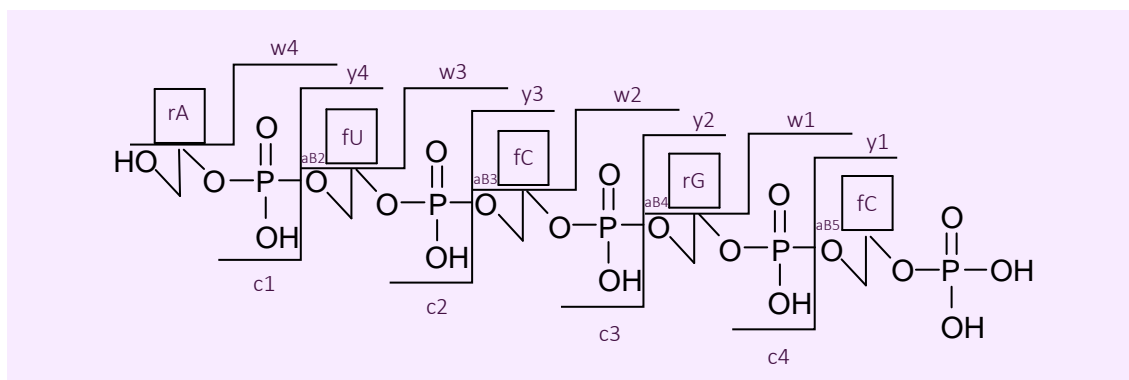


Figure 5.96. Fragment 4 of MinE07-139 5'-3'.
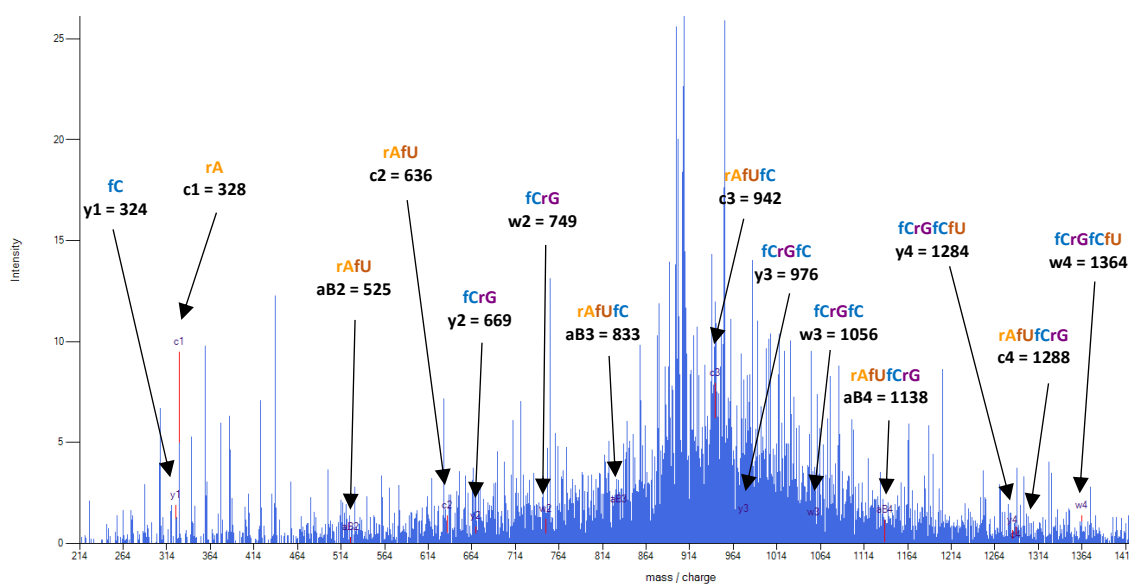


Figure 5.97. LC-MS/MS Fragment 4 Data of MinE07-139.

### 5.4.7.4.5. Fragment 5 of MinE07-139

5'-rGrGrAfU-VifUfUrArAfUfC **| rGfCfCrGfUrA |** rGrArArArArGfCrAfU-PhrGfU-VifCrArArArGfCfCr

GrGrArAfCfCrGfUfCfC-PCL-3'
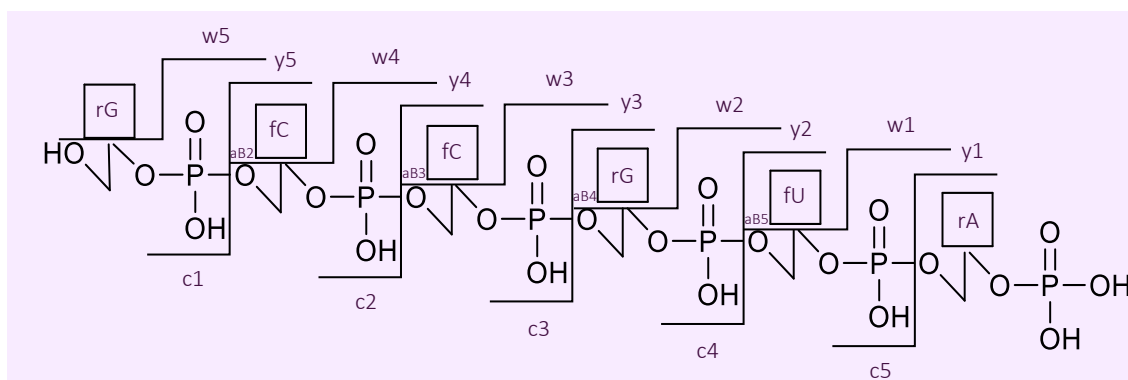


Figure 5.98. Fragment 5 of MinE07-139 5'-3'.
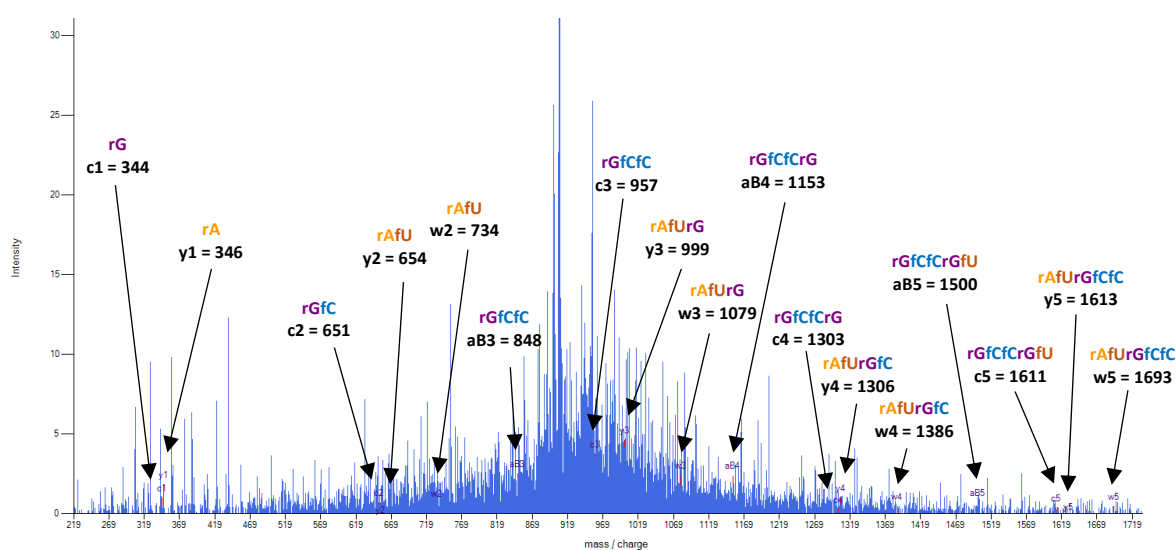


Figure 5.99. LC-MS/MS Fragment 5 Data of MinE07-139.

### 5.4.7.4.6. Fragment 6 of MinE07-139

5'-rGrGrAfU-VifUfUrArAfUfCrGfCfCrGfU **| rArGrArArArA |** rGfCrAfU-PhrGfU-VifCrArArArGfCfC

rGrGrArAfCfCrGfUfCfC-PCL-3'



Figure 5.100. Fragment 6 of MinE07-139 5'-3'.
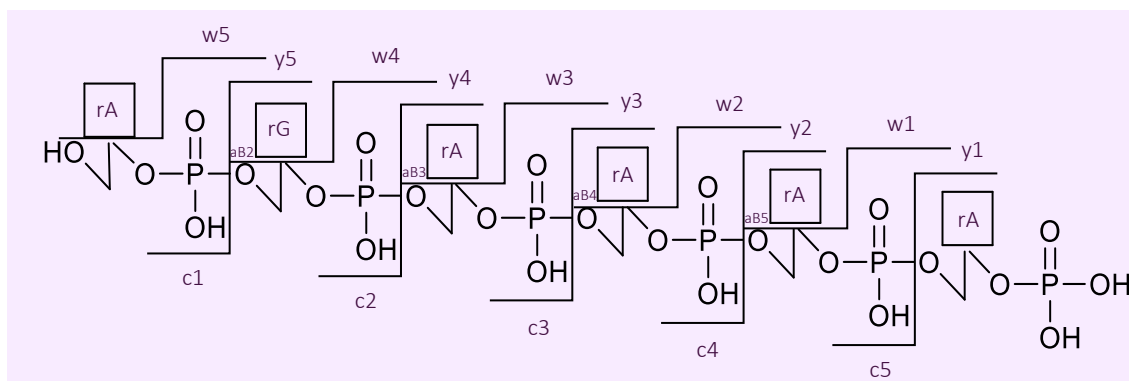


Figure 5.101. LC-MS/MS Fragment 6 Data of MinE07-139.

### 5.4.7.4.7. Fragment 7 of MinE07-139

5'-rGrGrAfU-VifUfUrArAfUfCrGfCfCrGfUrArGrArArArArGfC | **rAfU-PhrGfU-VifCrA** | rArArGfCfC

rGrGrArAfCfCrGfUfCfC-PCL-3'
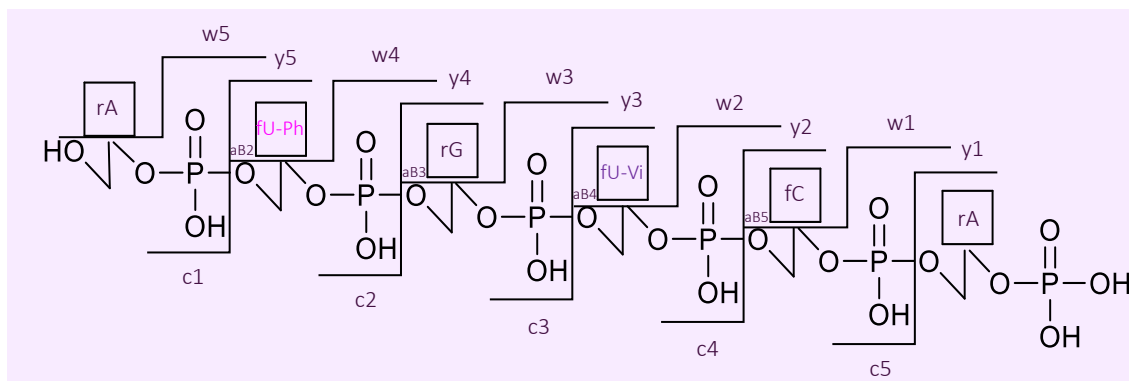


Figure 5.102. Fragment 7 of MinE07-139 5'-3'.
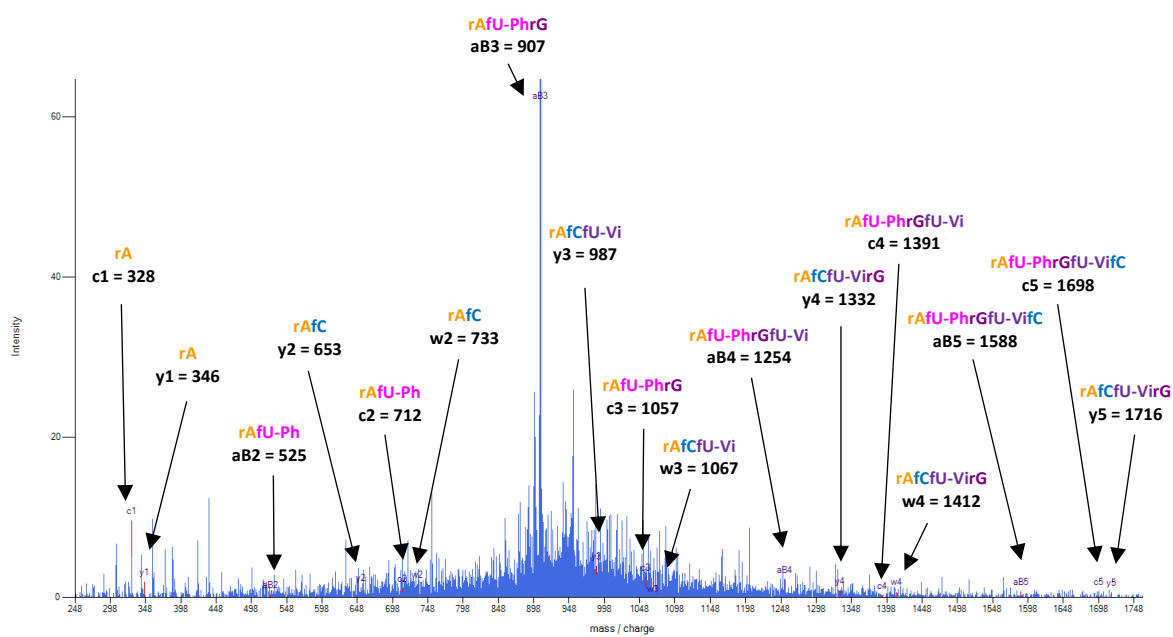


Figure 5.103. LC-MS/MS Fragment 7 Data of MinE07-139.

5'-rGrGrAfU-VifUfUrArAfUfCrGfCfCrGfUrArGrArArArArGfCrAfU-PhrGfU-Vi **| fCrArArArG |** fCfCr

GrGrArAfCfCrGfUfCfC-PCL-3'
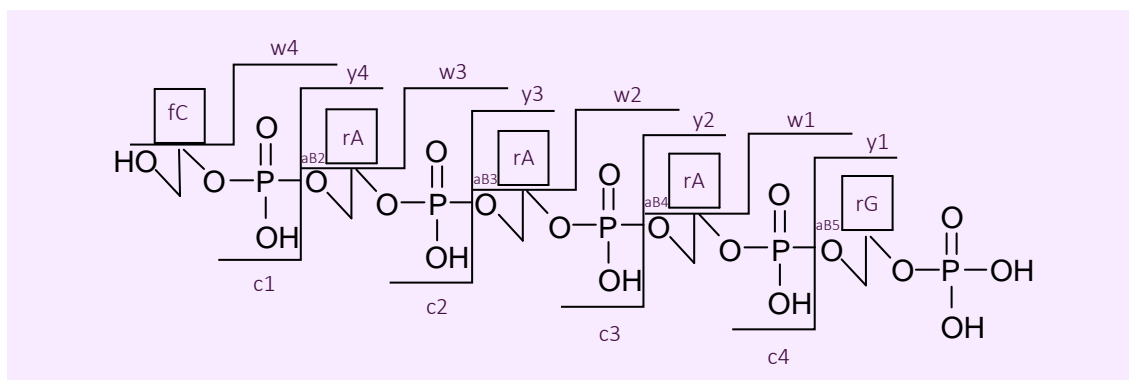


Figure 5.104. Fragment 8 of MinE07-139 5'-3'.
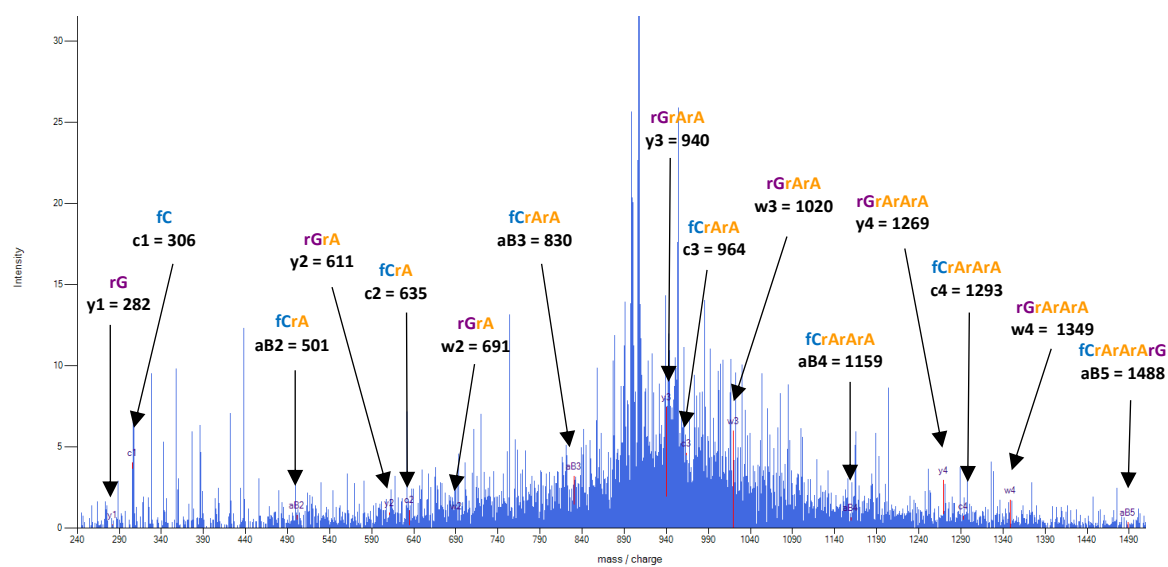


Figure 5.105. LC-MS/MS Fragment 8 Data of MinE07-139.

5'-rGrGrAfU-VifUfUrArAfUfCrGfCfCrGfUrArGrArArArArGfCrAfU-PhrGfU-VifCrArArArG **| fCfCrGr**

**GrArA |** fCfCrGfUfCfC-PCL-3'
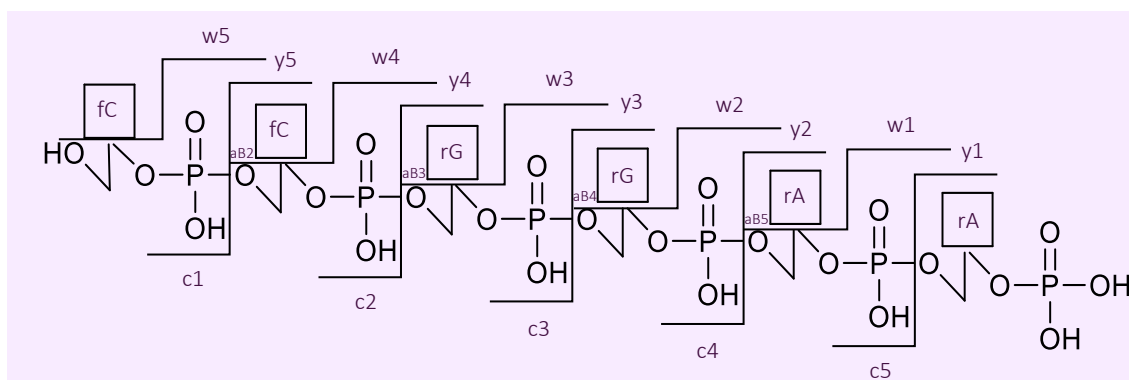


Figure 5.106. Fragment 9 of MinE07-139 5'-3'.
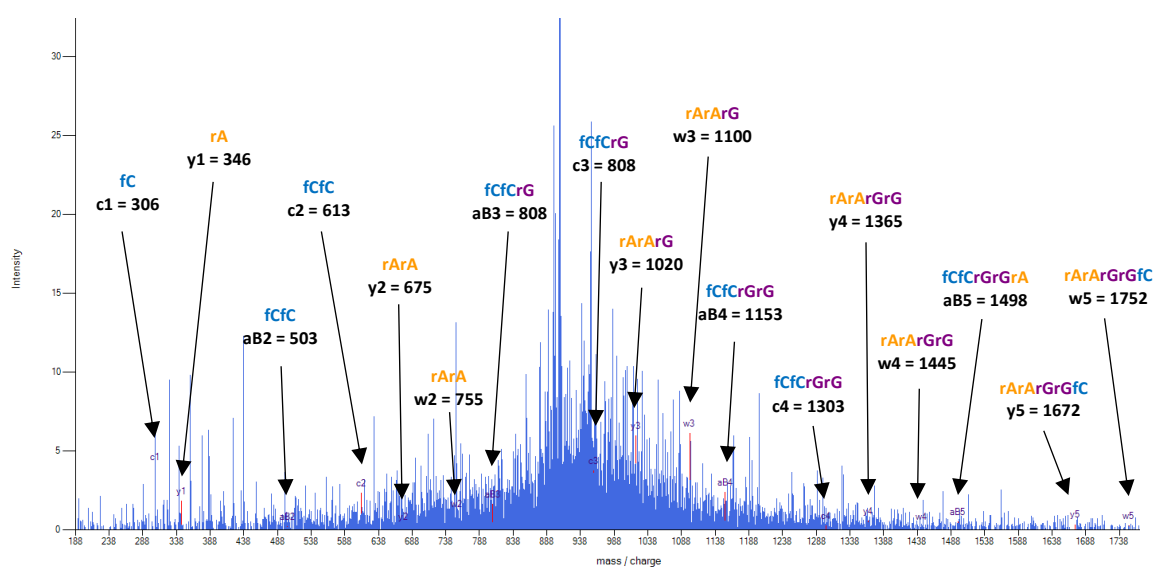


Figure 5.107. LC-MS/MS Fragment 9 Data of MinE07-139.

5'-rGrGrAfU-VifUfUrArAfUfCrGfCfCrGfUrArGrArArArArGfCrAfU-PhrGfU-VifCrArArArGfCfCrG **| r**

**GrArAfCfC |** rGfUfCfC-PCL-3'
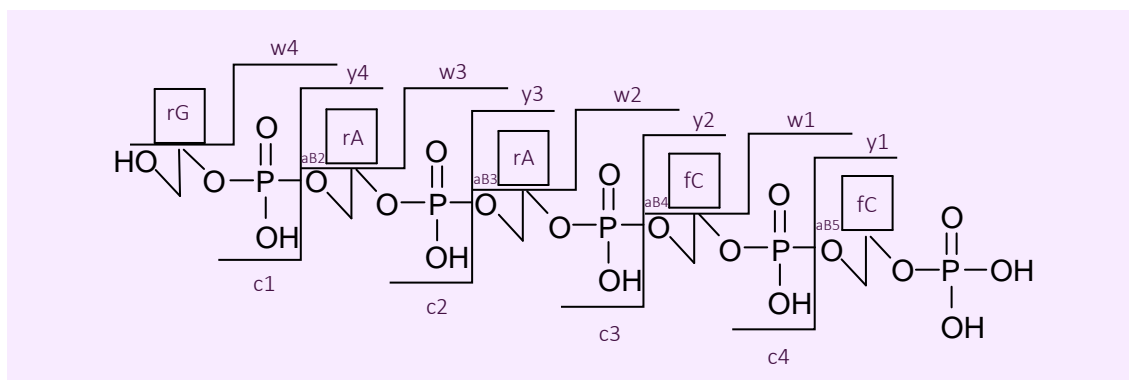


Figure 5.108. Fragment 10 of MinE07-139 5'-3'.
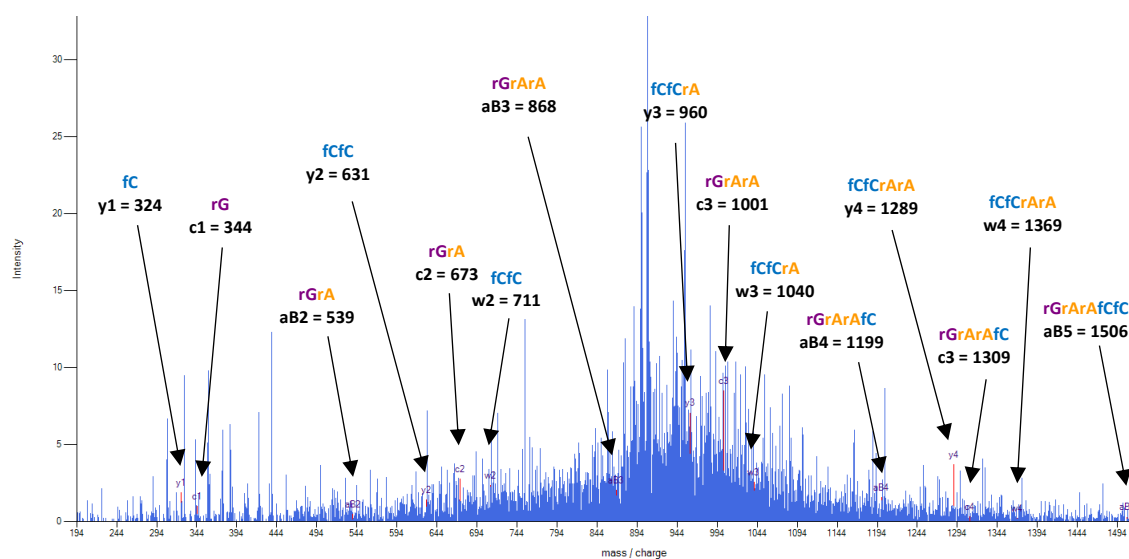


Figure 5.109. LC-MS/MS Fragment 10 Data of MinE07-139.

5'-rGrGrAfU-VifUfUrArAfUfCrGfCfCrGfUrArGrArArArArGfCrAfU-PhrGfU-VifCrArArArGfCfCrGrGr

ArAfC **| fCrGfUfCfC-PCL |** -3'
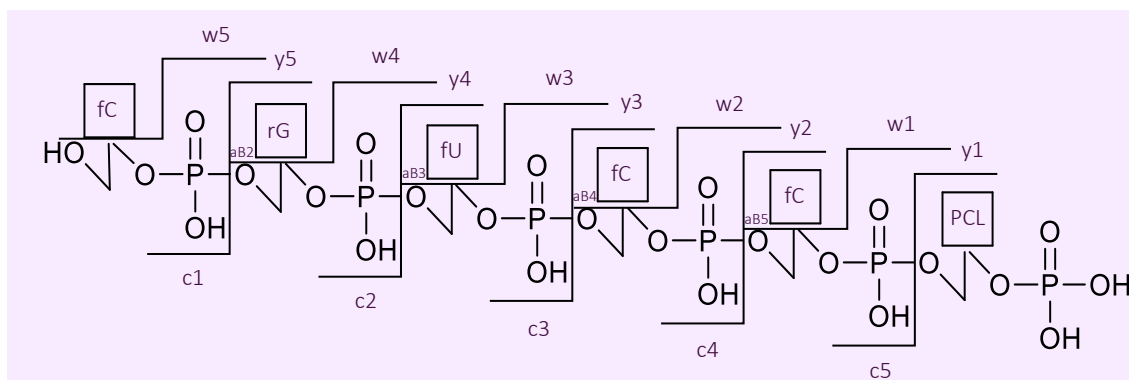


Figure 5.110. Fragment 11 of MinE07-139 5'-3'.
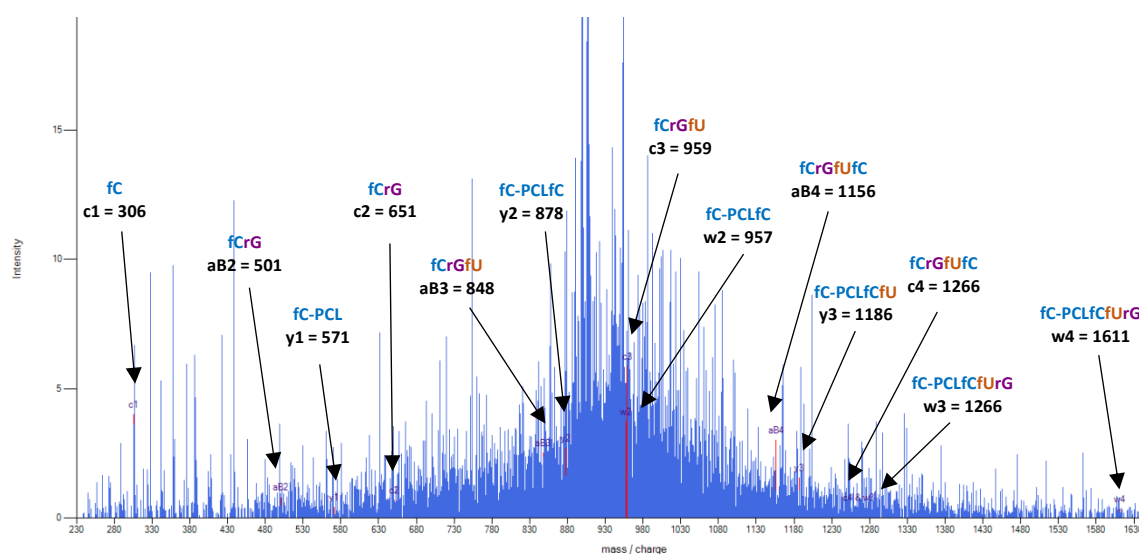


Figure 5.111. LC-MS/MS Fragment 11 Data of MinE07-139.

The data collected from **MinE07-139** is much more detailed than **MinE07-41** and **MinE07-56**.

The accuracy of the sequence will be confirmed by the protein binding affinity assays (chapter

6).

Figure 5.113 shows the mass spectrum of **MinE07-43** at RT of 39.185 minutes. This sample is below the MS/MS threshold and so there was no fragment pattern available for analysis.
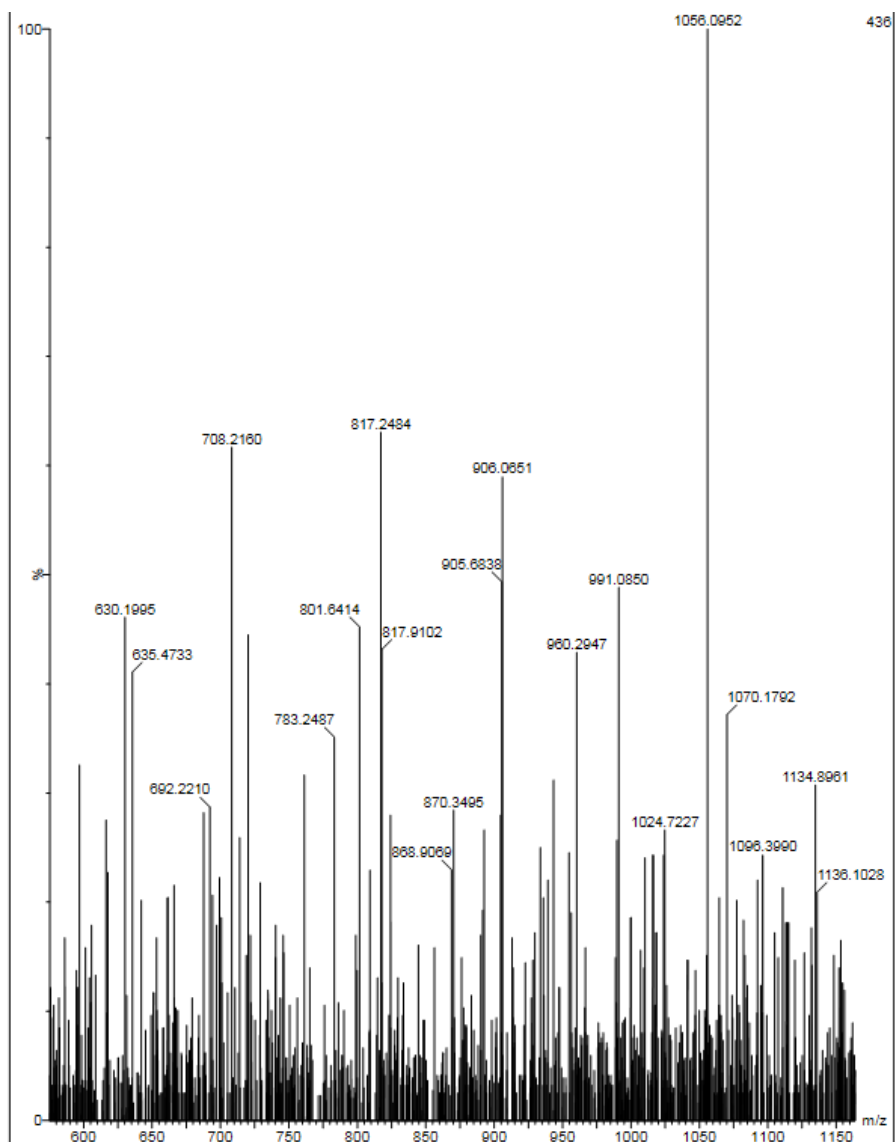


Figure 5.112. LC-MS spectra of MinE07-43.

## 5.5. Conclusion

In conclusion, the aptamer library was sorted successfully after some method development. 170 top aptamers were sorted into single wells before being photocleaved from their TentaGel microspheres solid support. Fifteen aptamers were run on the LC-MS/MS, three of which gave data to be analysed. They were successfully analysed using RoboOligo. These 3 aptamers **MinE07-41, MinE07-56** and **MinE07-139** will be re-synthesised for validation studies. The sequences **MinE07-41** and **MinE07-56** are relatively similar to each other. They both have three fU-Ph modified uridines, two of which are in the same location. They also both have at least one fU-I modified uridine. Potentially the two fU-Ph modified uridines in the same places are an important modification to MinE07 to help with affinity. Some molecular docking studies would help identify if this were true. **MinE07-139** also has a fU-Ph modified uridine, however not in the same locations as **MinE07-41** and **MinE07-56**. This chapter has confirmed that aptamers can be selected against their targets and selected out using flow cytometry and then identified using LC-MS/MS. Having a higher concentration of aptamer would have helped with their identification, improving the aptamers library synthesis may help with this in the future.

# Chapter 6

Biophysical and Biological Testing on Hit MinE07 Aptamer Sequences

# Chapter 6

## 6.1. Introduction

### 6.1.1. Protein Binding Affinity Assays

#### 6.1.1.2. Binding Affinity

Binding affinity is the strength of the binding interaction between a single biomolecule and its ligand. Binding affinity of proteins is measured by the equilibrium dissociation constant ($K_d$). The smaller the $K_d$ value the greater the binding affinity of the target to its ligand. The binding affinity is influenced by non-covalent intermolecular interactions such as hydrogen bonding, hydrophobic interactions, electrostatic interaction and Van der Waals forces between the target and ligand. Other molecules present can also effect the binding affinity by interfering or blocking the interactions.[506]

The aim of binding assays is to measure the interactions between two molecules. This includes proteins binding to small molecules, nucleic acids or even to another protein. The criteria for a successful binding assay to measure the affinity of a binding interaction between two molecules are that at the time of measurement the reaction must be at equilibrium, and for one reactant the concentration must vary.[507] The reactions are typically limited by the rates of collision between molecules which are determined by both the rates of diffusion of the molecules and the sizes of their interaction surfaces.[252]

$$Kd = \frac{K_-}{K_+} = [Aeq][Beq]/[ABeq]$$

The lower the value of $K_d$ the stronger the reaction (the more A and B are converted into AB). Given the small range of association rate constants, the value of the dissociation rate constant

often determines the affinity.[507] The equilibrium dissociation constant measures the tendency for the bound drug/target to dissociate to free drug and target. This is related to kinetic rate constants by the relationship $K_I = k_{off}/k_{on}$.[508]

## 6.1.2. Chemiluminescence Assays

Chemiluminescence is the luminescence produced by chemical reactions that induce the transition of an electron from its ground state to an excited electronic state. When the excited molecule then decays to the electronic ground state the chemiluminescence emission will occur at different wavelengths this can be from ultraviolet to visible to infrared radiation.[509]

Compared with fluorescence and absorbance assays, chemiluminescence assays have lower background signal and this equates to higher sensitivity. However, a lot of chemiluminescence based reactions suffer from low quantum efficiency and so produce weak luminescence. This potentially can restrict their applications in some analytical assays. In the past years, the development in nanomaterials has brough new capabilities into chemiluminescence assays. This is because of their great electronic, catalytic and optical properties, which can be used to improve the effectiveness of chemiluminescence assays.[510] Chemiluminescence is the method of signal detection in the binding assay.

## 6.1.3. Chapter Aims

The aim of this final chapter is resynthesis of the top hit aptamers and to evaluate their affinity for EGFR in a binding assay. A protein binding affinity assay has been run on **MinE07-Biotin-Cent, MinE07-U-Ph-Biotin, MinE07-U-Vi-Biotin, MinE07-U-I-Biotin, MinE07-41-Biotin, MinE07-41.2-Biotin** (non-truncated version of MinE07-41), **MinE07-56-Biotin, MinE07-139-Biotin** and **MinE07-139.2-Biotin** (non-truncated version of MinE07-139),. These protein binding affinity assays will be monitored by level of chemiluminescence. The assay has been run on protein A plates, EGFR will have an Fc tag of which has a high affinity for protein A. The aptamers

will have a biotin tag so it can bind with streptavidin-AP followed by tropix CPD star chemiluminescence substrate. This will be done to generate $K_d$ values take can be compared. These results will show whether this novel strategy for the synthetic selection of enhanced therapeutic aptamers is successful.

# 6.2. Materials and Instrumentation

## 6.2.1. Materials

Synthesiser reagents: oxidizer (0.02M iodine, 20% pyridine), Cap A Mix (THF/Pyridine/acetic anhydride 8:1:1), Cap B Mix (10% methylimidazole in THF), de-block (3% trimethylamine in DCM) and ETT activator (0.25 M, 5-ethylthio-1H-tetrazole in acetonitrile) were purchased from Link Technologies. N2-acetyl-2'-O-tert-butyldimethylsilyl-5'-O-DMT-guanosine 3'-CE phosphoramidite, N6-benzoyl-2'-O-tert-butyldimethylsilyl-5'-O-DMT-adenosine 3'-CE phosphoramidite, N4-acetyl-2'-deoxy-5'-O-DMT-2'-fluorocytidine 3'-CE phosphoramidite and 2'-Deoxy-5'-O-DMT-2'-fluorouridine 3'-CE phosphoramidite were purchased from CarboSynth. Sheath Fluid was purchased from BD Biosciences. Recombinant Human EGFR Fc Chimera Protein, CF was purchased from R&D Systems. Native *Staphylococcus aureus Protein A* (FITC) (ab7455) was purchased from Abcam. Bovine Serum Albumin, Nuclease Free Water, Salmon Sperm DNA (ssDNA), PBS with calcium and magnesium (PBS$^{(+/+)}$), PBS without calcium and magnesium (PBS$^{(-/-)}$), Tween-20 and Streptavidin–Alkaline Phosphatase (Streptavidin–AP) from *Streptomyces avidinii* were purchased from Sigma Aldrich. Tropix CDP-Star Ready to use with Sapphire II was purchased from AB Biosystem. White protein A plates were purchased from Biomat. Biotin-MINE07 was purchased from Integrated DNA Technologies.

### 6.2.1.1. Buffers

Wash Buffer: PBS + 0.05% Tween 20, to 500 mL of PBS $^{(+/+)}$, add 250 μL of Tween-20.

Protein Binding Buffer: 48.5 mL of PBS $^{(+/+)}$, 1 mL 50 mg/ml BSA and 0.5 mL 10 mg/mL ssDNA were added

## 6.2.2. Instrumentation

### 6.2.2.1. Expedite™ 8909 DNA Synthesiser

Specification can be found in Chapter 2 section 2.2.2.1 Expedite™ 8909 DNA Synthesiser

### 6.2.2.2. Mass Spectrometer

Specification can be found in Chapter 3 section 3.2.2.3 Mass Spectrometer

### 6.2.2.1. Plater Reader

Specification can be found in Chapter 5 section 5.2.2.3 Plater Reader

### 6.2.2.3. UV-Visible Absorption

Specification can be found in Chapter 2 section 2.2.2.5 UV-Visible Absorption Spectra

# 6.3. Experimental

## 6.3.1. Synthesis of MinE07-41-Biotin Aptamer

Universal UnyLinker support (0.021 g) was added to a synthesiser column. See experimental method 4.3.1 for the setup of the synthesiser. The sequence uploaded on to the synthesiser was 5'-Biotin-fCrGrGrAfUfU-PhfU-PhrArAfUfCrGfCfCrGfU-IrArGrArArArArGfCrAfUrGfUfCrArAr ArGfCfCrGrGrArAfCfCrGfU-PhfCfC-3'. This was run to completion (28.9% yield).

## 6.3.2. Synthesis of MinE07-41.2-Biotin Aptamer

Universal UnyLinker support (0.021 g) was added to a synthesiser column. See experimental method 4.3.1 for the setup of the synthesiser. The sequence uploaded on to the synthesiser

was 5'-Biotin-rGrGrAfCrGrGrAfUfU-PhfU-PhrArAfUfCrGfCfCrGfU-IrArGrArArArArGfCrAfUrGfUf CrArArArGfCfCrGrGrArAfCfCrGfU-PhfCfC-3'. This was run to completion (15.8% yield).

### 6.3.3. Synthesis of MinE07-56-Biotin Aptamer

Universal UnyLinker support (0.021 g) was added to a synthesiser column. See experimental method 4.3.1 for the setup of the synthesiser. The sequence uploaded on to the synthesiser was 5'-Biotin-rGrGrAfCrGrGrAfUfU-PhfU-PhrArAfU-IfCrGfCfCrGfU-PhrArGrArArArArGfCrAfUrGf UfCrArArArGfCfCrGrGrArAfCfCrGfUfCfC-3'. This was run to completion (21.4% yield).

### 6.3.4. Synthesis of MinE07-139-Biotin Aptamer

Universal UnyLinker support (0.021 g) was added to a synthesiser column. See experimental method 4.3.1 for the setup of the synthesiser. The sequence uploaded on to the synthesiser was 5'-Biotin-rGrGrAfU-VifUfUrArAfUfCrGfCfCrGfUrArGrArArArArGfCrAfU-PhrGfU-VifCrArArAr GfCfCrGrGrArAfCfCrGfUfCfC-3'. This was run to completion (18.4% yield).

### 6.3.5. Synthesis of MinE07-139.2-Biotin Aptamer

Universal UnyLinker support (0.021 g) was added to a synthesiser column. See experimental method 4.3.1 for the setup of the synthesiser. The sequence uploaded on to the synthesiser was 5'-Biotin-rGrGrAfCrGrGrAfU-VifUfUrArAfUfCrGfCfCrGfUrArGrArArArArGfCrAfU-PhrGfU-Vif CrArArArGfCfCrGrGrArAfCfCrGfUfCfC-3'. This was run to completion (21.4% yield).

### 6.3.6. Cleavage of MinE07-41-Biotin, MinE07-41.2-Biotin, MinE07-56-Biotin, MinE07-139-Biotin and MinE07-139.2-Biotin Aptamers from Universal Unylinker Support

MinE07-41-Biotin, MinE07-41.2-Biotin, MinE07-56-Biotin, MinE07-139-Biotin and MinE07-139.2-Biotin aptamers were cleaved and deprotected following the same protocol as seen in 4.3.20.

### 6.3.7. Desalt of MinE07-41-Biotin, MinE07-41.2-Biotin, MinE07-56-Biotin, MinE07-139-Biotin and MinE07-139.2-Biotin Aptamers using Zetadex

MinE07-41-Biotin, MinE07-41.2-Biotin, MinE07-56-Biotin, MinE07-139-Biotin and MinE07-139.2-Biotin aptamers were desalted following the same protocol as seen in 4.3.21.

### 6.3.8. MinE07-41-Biotin, MinE07-41.2-Biotin, MinE07-56-Biotin, MinE07-139-Biotin and MinE07-139.2-Biotin Aptamers on 1.5 mm 15% Denaturing PAGE Purification Gel

MinE07-41-Biotin, MinE07-41.2-Biotin, MinE07-56-Biotin, MinE07-139-Biotin and MinE07-139.2-Biotin aptamers were purified following the same protocol as seen in 4.3.23.

### 6.3.9. Ethanol Precipitation MinE07-41-Biotin, MinE07-41.2-Biotin, MinE07-56-Biotin, MinE07-139-Biotin and MinE07-139.2-Biotin after PAGE Purification

MinE07-41-Biotin, MinE07-41.2-Biotin, MinE07-56-Biotin, MinE07-139-Biotin and MinE07-139.2-Biotin aptamers were ethanol precipitated following the same protocol as seen in 4.3.24.

### 6.3.10. Analysis of MinE07-41-Biotin, MinE07-41.2-Biotin, MinE07-56-Biotin, MinE07-139-Biotin and MinE07-139.2-Biotin Concentrations.

MinE07-41-Biotin, MinE07-41.2-Biotin, MinE07-56-Biotin, MinE07-139-Biotin and MinE07-139.2-Biotin aptamers were analysised following the same protocol as seen in 4.3.25.

### 6.3.11. HABA Biotin Binding Assay with MinE07-Biotin-Cent Aptamer.

Streptavidin-AP at a 1:1000 dilution in binding buffer was prepared, 20 µL was added to each well and it was incubated for 30 minutes at room temperature shaking at 450 rpm with 20 µl of Tropix CDP-Star ready to use substrate. The plate was read using the chemiluminescence protocol on the Victor X4 plate reader to check it was giving a signal. 20 µL of HABA (5 µM) was added to the wells and it was incubated for 30 minutes at room temperature shaking at 450 rpm. MinE07-Biotin-Cent was added to relevant wells and incubated for 30 minutes at room

temperature shaking at 450 rpm. The plate was then read again using the chemiluminescence

protocol was then read on the Victor X4 plate reader.

### 6.3.12. HABA Biotin Binding Assay with MinE07-41-Biotin, MinE07-41.2-Biotin, MinE07-56-Biotin, MinE07-139-Biotin and MinE07-139.2-Biotin Aptamers.

Streptavidin-AP at a 1:1000 dilution in binding buffer was prepared, 20 µL was added to each

well and it was incubated for 30 minutes at room temperature shaking at 450 rpm with 20 µl of

Tropix CDP-Star ready to use substrate. The plate was then read using the chemiluminescence

protocol on the Victor X4 plate reader to check it was giving a signal. 20 µL of HABA (5 µM) was

added to the wells and it was incubated for 30 minutes at room temperature shaking at 450rpm.

MinE07-41-Biotin, MinE07-41.2-Biotin, MinE07-56-Biotin, MinE07-139-Biotin and MinE07-

139.2-Biotin were added to relevant wells and incubated for 30 minutes at room temperature

shaking at 450rpm. The plate was then read using the chemiluminescence protocol on the Victor

X4 plate reader.

### 6.3.13. MinE07-41-Biotin, MinE07-41.2-Biotin, MinE07-56-Biotin, MinE07-139-Biotin and MinE07-139.2-Biotin Aptamers on 15% Denaturing Polyacrylamide Gel

MinE07-41-Biotin, MinE07-41.2-Biotin, MinE07-56-Biotin, MinE07-139-Biotin and MinE07-

139.2-Biotin aptamers were run on an analytical gel following the same protocol as seen in

4.3.22.

### 6.3.14. Protein Binding Affinity Assay of MinE07M3-Bio and MinE07M6-Bio compared with parent aptamer MinE07-Bio-Cent

The protein A plates were washed 3 times with 150 µL of wash buffer per well. EGFR-Fc protein

was prepared in wash buffer at 1 µg/mL. 100 µL EGFR-Fc coating buffer per well was added and

incubated for 30 minutes at 450 rpm at room temperature. The liquid was flicked out of the

well plate. The plate was washed 3 times with wash buffer and then 50 µL 1x binding buffer to

each well was added, and then incubated at 450 rpm for 10 minutes. MinE07-Bio-Cent/MinE07M3-Bio/MinE07M6-Bio aptamers were prepared with a final concentration of 10 µM in PBS(-/-). The diluted aptamer was then denatured in the Eppendorf master cycler PCR machine (85°C for 5 minutes, then cooled to 25°C at 0.1°C per second and held at 25°C). The binding buffer was removed from the plate, which was then washed 3 times with wash buffer. 90 µL of binding buffer was added to each well followed by 10 µL of aptamer to the relevant wells. This was then incubated for 60 minutes at room temperature at 450rpm. The liquid was flicked out of the well plate, and it was washed 3 times with wash buffer. Streptavidin-AP at a 1:1000 dilution in binding buffer was prepared, 100 µL was added to each well and it was incubated for 30 minutes at room temperature shaking at 450 rpm. The plate was then washed with wash buffer and then 95 µL of Tropix CDP-Star ready to use substrate was added and incubated for 10 minutes with shaking at 450 rpm. The chemiluminescence protocol was used to read the plate on a plate reader.

## 6.3.15. Protein Binding Affinity Assay of MinE07-41-Biotin, MinE07-41.2-Biotin, MinE07-56-Biotin, MinE07-139-Biotin and MinE07-139.2-Biotin

The protein A plates were washed 3 times with 150 µL of wash buffer per well. EGFR-Fc protein was prepared in wash buffer at 1 µg/mL. 100 µl EGFR-Fc coating buffer per well was added and incubated for 30 minutes at 450 rpm at room temperature. The liquid was flicked out of the well plate. The plate was washed 3 times with wash buffer and then 50 µL 1x binding buffer to each well was added, and then incubate at 450 rpm for 10 minutes. MinE07-41-Biotin, MinE07-41.2-Biotin, MinE07-56-Biotin, MinE07-139-Biotin and MinE07-139.2-Biotin aptamers were prepared at concentrations 500 nM, 158 nM, 50 nM, 15.8 nM, 5 nM, 1.58 nM, 1 nM and 0.5 nM in binding buffer. The diluted aptamer was then denatured in the Eppendorf master cycler PCR machine (85°C for 5 minutes, then cooled to 25°C at 0.1°C per second and held at 25°C). The binding buffer was removed from the plate and washed 3 times with wash buffer. 100 µL of

aptamer was added to the relevant wells. This was then incubated for 60 minutes at room temperature at 450 rpm. The liquid was flicked out of the well plate, and it was washed 3 times with wash buffer. Streptavidin-AP at a 1:1000 dilution in binding buffer was prepared, 100 μL was added to each well and it was incubated for 30 minutes at room temperature shaking at 450 rpm. The plate was then washed with wash buffer and then 95 μL of Tropix CDP-Star ready to use substrate was added and incubated for 10 minutes with shaking at 450rpm. The plate was then read again using the chemiluminescence protocol was then read on the Victor X4 plate reader. This was repeated until 3 biological repeats were achieved.

### 6.3.16. Data Analysis of the Protein Binding Affinity Assay Results

Data analysis was done on Prism GraphPad Version 8 using the one site - specific binding model to calculated Bmax and $K_d$.

# 6.4. Results and Discussion

## 6.4.1. Re-synthesis of Hit Aptamers MinE07-41-Biotin, MinE07-41.2-Biotin, MinE07-56-Biotin, MinE07-139-Biotin and MinE07-139.2-Biotin

### 6.4.1.1. Re-synthesis of Hit Aptamers MinE07-41-Biotin

MinE07-41-Biotin: 5'-Biotin-fCrGrGrGrAfUfU-PhfU-PhrArAfUfCrGfCfCrGfU-IrArGrArArArArArGfCrAf UrGfUfCrArArArArGfCfCrGrGrArAfCfCrGfU-PhfCfC-3'.

**MinE07-41-Biotin** is re-synthesised so it can be validated as the correct sequence and having a higher binding affinity for EGFR. This synthesis was successfully shown in the trityl monitor for this aptamer (figure 6.1). The trityl monitor gave a good coupling efficiently of $1.46e^{6}$.
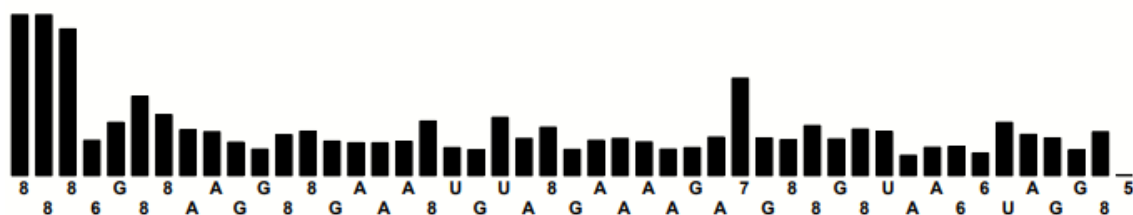
Figure 6.1. Trityl Monitor of the Synthesis of MinE07-41-Biotin on Universal UnyLinker Solid Support Cytosine is labelled as 8, fU-Ph is labelled at 6, fU-I is labelled as 7 and the biotin is labelled as 5.

The step wise coupling efficiency is 97.5% as shown in the synthesis report (figure 6.2). To get the overall average yield: *0.975$^{49}$ = 0.289.* This gives a 28.9% overall synthetic yield. This yield is good enough to carry forward for purification, HABA assay studies and binding affinity studies.
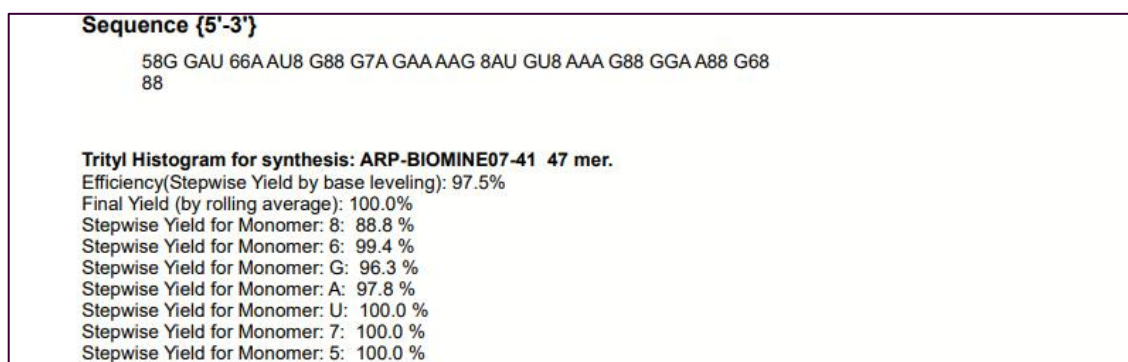


Figure 6.2. Aptamer Synthesis Report of the Synthesis of MinE07-41-Biotin on Universal UnyLinker Solid Support.

### 6.4.1.2. Re-synthesis of Hit Aptamer MinE07-41.2-Biotin

MinE07-41.2-Biotin: 5'-Biotin-rGrGrAfCrGrGrAfUfU-PhfU-PhrArAfUfCrGfCfCrGfU-IrArGrArArAr ArGfCrAfUrGfUfCrArArArArGfCfCrGrGrArAfCfCrGfU-PhfCfC-3'.

**MinE07-41.2-Biotin** is re-synthesised so it can be validated as the correct sequence and having a higher binding affinity for EGFR. This synthesis was successfully shown in the trityl monitor for this aptamer (figure 6.3). The trityl monitor gave a good coupling efficiently of 1.44e$^6$.
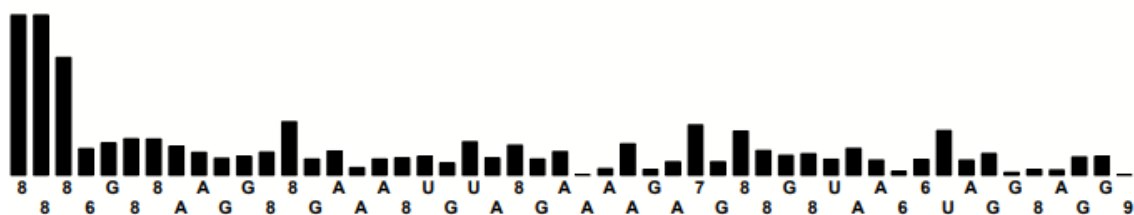
Figure 6.3. Trityl Monitor of the Synthesis of MinE07-41.2-Biotin on Universal UnyLinker Solid Support Cytosine is labelled as 8, fU-Ph is labelled at 6, fU-I is labelled as 7 and the biotin is labelled as 9.

The step wise coupling efficiency is 96.3% as shown in the synthesis report (figure 6.4). To get the overall average yield: *0.963[49] = 0.158.* This gives a 15.8% overall synthetic yield. This yield is good enough to carry forward for purification, HABA assay studies and binding affinity studies.

### Sequence {5'-3'}

9GG A8G GAU 66A AU8 G88 G7A GAA AAG 8AU GU8 AAA G88 GGA A88 G68 88

**Trityl Histogram for synthesis: ARP-BIOMINE07-412  50 mer.**
Efficiency(Stepwise Yield by base leveling): 96.3%
Final Yield (by rolling average): 100.0%
Stepwise Yield for Monomer: 8:  89.2 %
Stepwise Yield for Monomer: 6:  97.8 %
Stepwise Yield for Monomer: G:  95.4 %
Stepwise Yield for Monomer: A:  91.9 %
Stepwise Yield for Monomer: U:  100.0 %
Stepwise Yield for Monomer: 7:  100.0 %
Stepwise Yield for Monomer: 9:  100.0 %
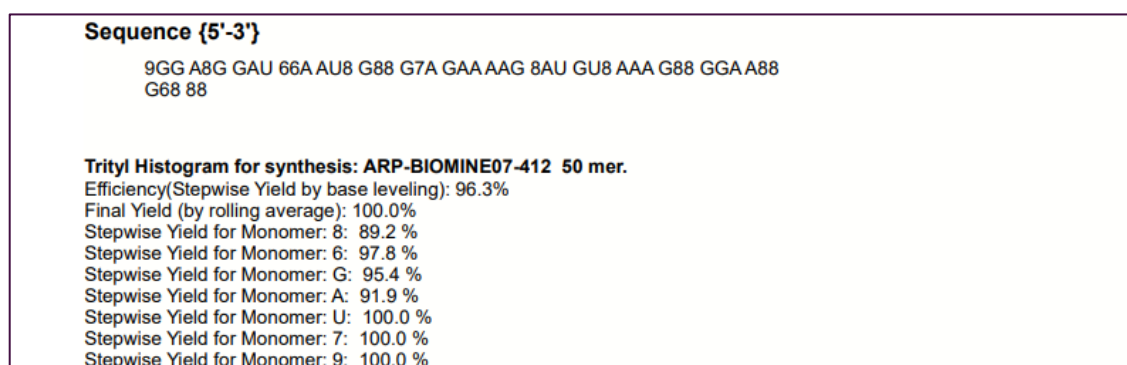
Figure 6.4. Aptamer Synthesis Report of the Synthesis of MinE07-41.2-Biotin on Universal UnyLinker Solid Support.

Both **MinE07-41-Biotin** and **MinE07-41.2-Biotin** were cleaved from their solid support and purified by PAGE purification. The samples were then desalted with Zetadex resin and run on an analytical polyacrylamide gel to check that the molecular weight was correct for the length of this aptamer. Figure 6.5 below shows an analytical gel of **MinE07-41-Biotin** and **MinE07-41.2-Biotin** after page purification, showing they are pure enough to take forward for further analysis.
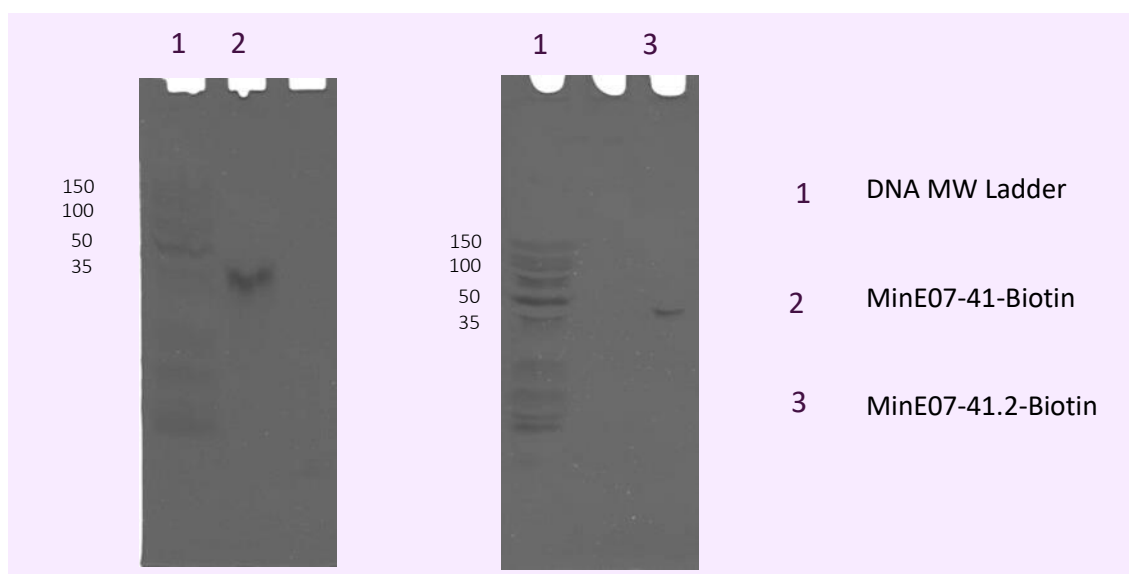
Figure 6.5. 15% Polyacrylamide Gel of MinE07-41-Biotin and MinE07-41.2-Biotin after PAGE purification.

### 6.4.1.3. Re-synthesis of Hit Aptamer MinE07-56-Biotin

MinE07-56-Biotin: 5'-Biotin-rGrGrGrAfCrGrGrGrAfUfU-PhfU-PhrArAfU-IfCrGfCfCrGfU-PhrArGrArArA rArGfCrAfUrGfUfCrArArArGfCfCrGrGrArAfCfCrGfUfCfC-3'

**MinE07-56-Biotin** is re-synthesised so it can be validated as the correct sequence and having a higher binding affinity for EGFR. This synthesis was successfully shown in the trityl monitor for this aptamer (figure 6.6). The trityl monitor gave a good coupling efficiently of $1.44e^{6}$.
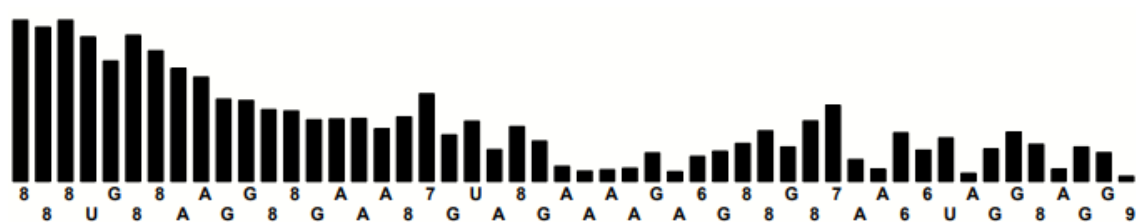


Figure 6.6. Trityl Monitor of the Synthesis of MinE07-56-Biotin on Universal UnyLinker Solid Support Cytosine is labelled as 8, fU-Ph is labelled at 6, fU-I is labelled as 7 and the biotin is labelled as 9.

The step wise coupling efficiency is 96.3% as shown in the synthesis report (figure 6.7). To get the overall average yield: *0.969$^{49}$ = 0.214.* This gives a 21.4% overall synthetic yield. This yield is good enough to carry forward for purification, HABA assay studies and binding affinity studies.



**Sequence {5'-3'}**
9GG A8G GAU 66A A78 G88 G6A GAA AAG 8AU G78 AAA G88 GGA A88
GU8 88

**Trityl Histogram for synthesis: ARP-BIOMINE07-56  50 mer.**
Efficiency(Stepwise Yield by base leveling): 96.9%
Final Yield (by rolling average): 100.0%
Stepwise Yield for Monomer: 8:  94.3 %
Stepwise Yield for Monomer: U:  97.3 %
Stepwise Yield for Monomer: G:  95.4 %
Stepwise Yield for Monomer: A:  92.0 %
Stepwise Yield for Monomer: 7:  99.6 %
Stepwise Yield for Monomer: 6:  100.0 %
Stepwise Yield for Monomer: 9:  100.0 %
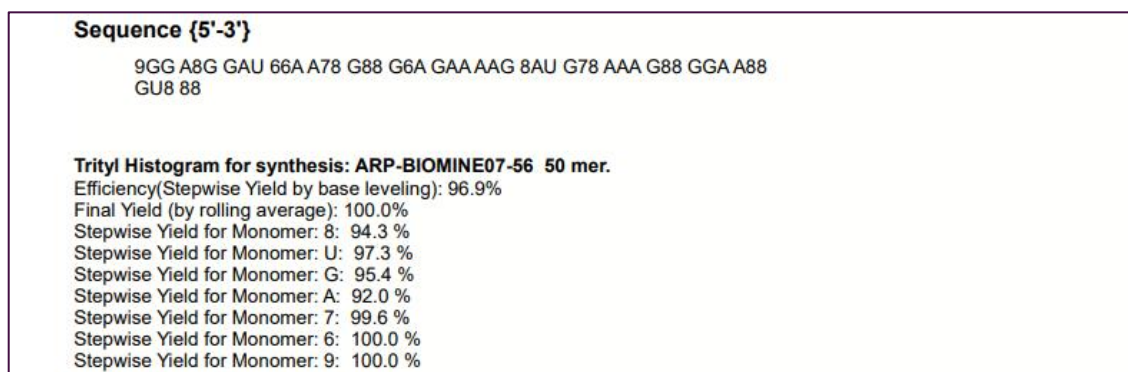
Figure 6.7. Aptamer Synthesis Report of the Synthesis of MinE07-56-Biotin on Universal UnyLinker Solid Support.

This sample was then cleaved from the solid support and purified by PAGE purification. The sample was then desalted with Zetadex resin and run on an analytical polyacrylamide gel showing the correct molecular weight was correct for the length of this aptamer (figure 6.8).



| | |
|---|---|
| 1 | DNA MW Ladder |
| 4 | MinE07-56-Biotin |
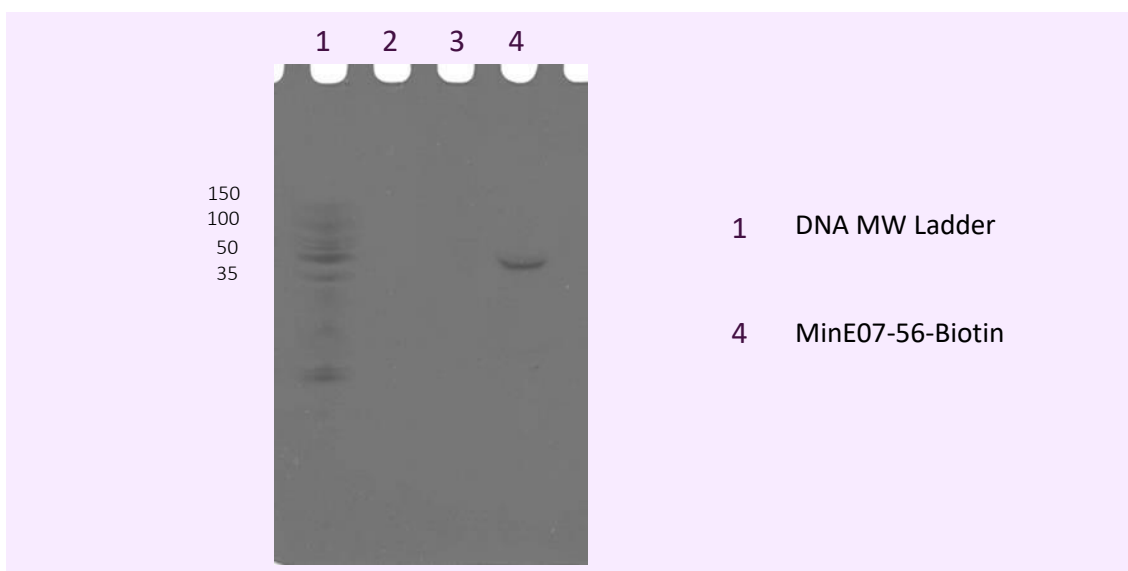
Figure 6.8. 15% Polyacrylamide Gel of MinE07-56-Biotin after PAGE purification.

6.4.1.4. Re-synthesis of Hit Aptamer MinE07-139-Biotin

MinE07-139-Biotin: 5'-Biotin-rGrGrAfU-VifUfUrArAfUfCrGfCfCrGfUrArGrArArArArGfCrAfU-Phr GfU-VifCrArArArGfCfCrGrGrArAfCfCrGfUfCfC-3'

**MinE07-139-Biotin** is re-synthesised so it can be validated as the correct sequence and having a higher binding affinity for EGFR. This synthesis was successfully shown in the trityl monitor for this aptamer (figure 6.9). The trityl monitor gave a good coupling efficiently of $1.47e^6$.
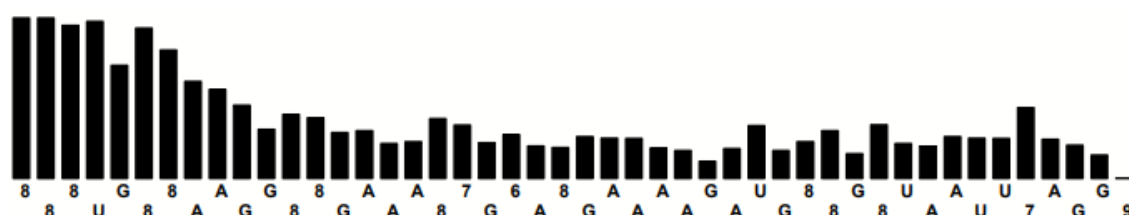


Figure 6.9. Trityl Monitor of the Synthesis of MinE07-139-Biotin on Universal UnyLinker Solid Support Cytosine is labelled as 8, fU-Ph is labelled at 6, fU-Vi is labelled as 7 and the biotin is labelled as 9.

The step wise coupling efficiency is 96.3% as shown in the synthesis report (figure 6.10). To get the overall average yield: *$0.966^{49}$ = 0.184.* This gives an 18.4% overall synthetic yield. This yield is good enough to carry forward for purification, HABA assay studies and binding affinity studies.



**Sequence {5'-3'}**

9GG A7U UAA U8G 88G UAG AAA AG8 A6G 78A AAG 88G GAA 88G U88 8

**Trityl Histogram for synthesis: ARP-BIOMINE07-139C8  46 mer.**
Efficiency(Stepwise Yield by base leveling): 96.6%
Final Yield (by rolling average): 100.0%
Stepwise Yield for Monomer: 8:  91.3 %
Stepwise Yield for Monomer: U:  96.9 %
Stepwise Yield for Monomer: G:  93.9 %
Stepwise Yield for Monomer: A:  94.1 %
Stepwise Yield for Monomer: 7:  100.0 %
Stepwise Yield for Monomer: 6:  100.0 %
Stepwise Yield for Monomer: 9:  100.0 %
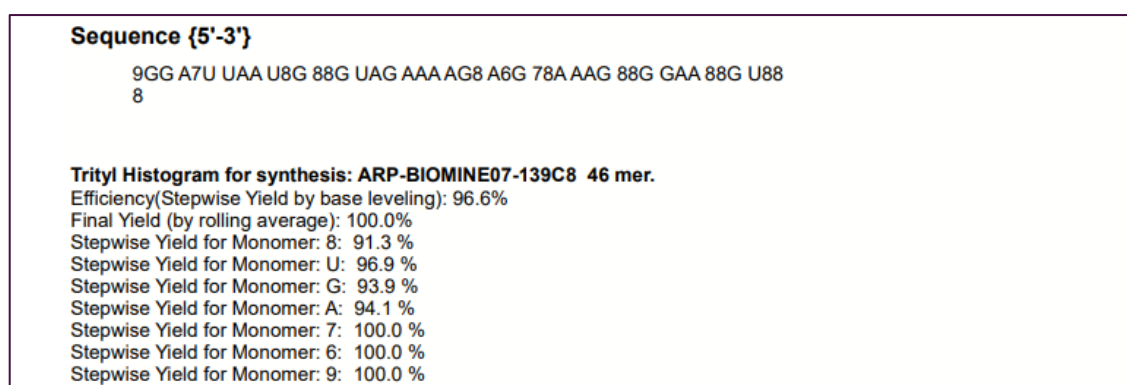
Figure 6.10. Aptamer Synthesis Report of the Synthesis of MinE07-139-Biotin on Universal UnyLinker Solid Support.

### 6.4.1.5. Re-synthesis of Hit Aptamer MinE07-139.2-Biotin

MinE07-139.2-Biotin: 5'-Biotin-rGrGrAfCrGrGrAfU-VifUfUrArAfUfCrGfCfCrGfUrArGrArArArArGf

CrAfU-PhrGfU-VifCrArArArGfCfCrGrGrArAfCfCrGfUfCfC-3'

**MinE07-139.2-Biotin** is re-synthesised so it can be validated as the correct sequence and having
a higher binding affinity for EGFR. This synthesis was successfully shown in the trityl monitor for
this aptamer (figure 6.11). The trityl monitor gave a good coupling efficiently of $1.38e^6$.
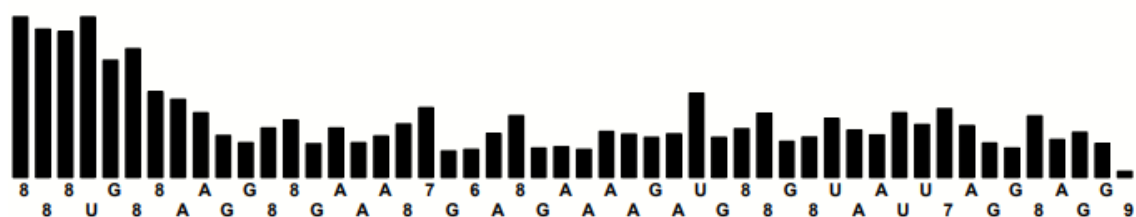


Figure 6.11. Trityl Monitor of the Synthesis of MinE07-139.2-Biotin on Universal UnyLinker Solid Support Cytosine is
labelled as 8, fU-Ph is labelled at 6, fU-Vi is labelled as 7 and the biotin is labelled as 9.

The step wise coupling efficiency is 96.3% as shown in the synthesis report (figure 6.12). To get
the overall average yield: *0.969<sup>49</sup> = 0.214.* This gives a 21.4% overall synthetic yield. This yield is
good enough to carry forward for purification, HABA assay studies and binding affinity studies.
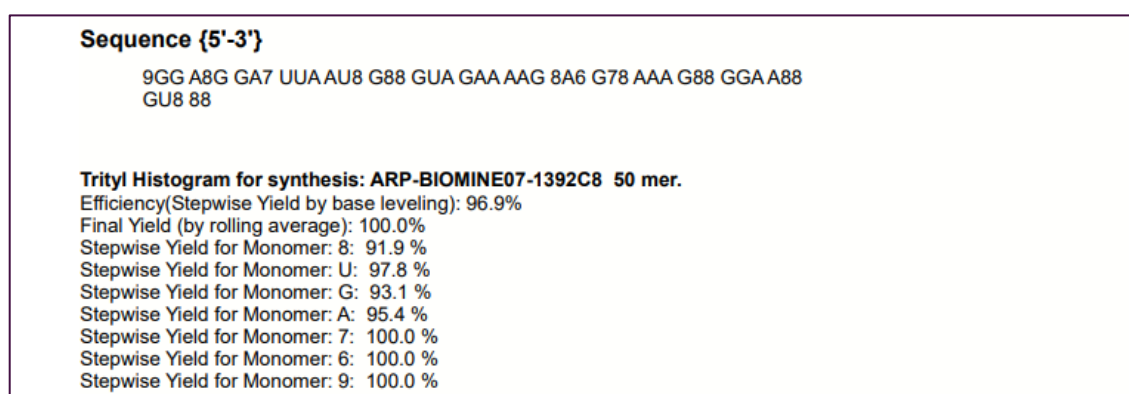


Figure 6.12. Aptamer Synthesis Report of the Synthesis of MinE07-139.2-Biotin on Universal UnyLinker Solid
Support.

This sample was then cleaved from the solid support and purified by PAGE purification. The sample was then desalted with Zetadex resin and run on an analytical polyacrylamide gel showing the correct molecular weight was correct for the length of this aptamer (figure 6.13).
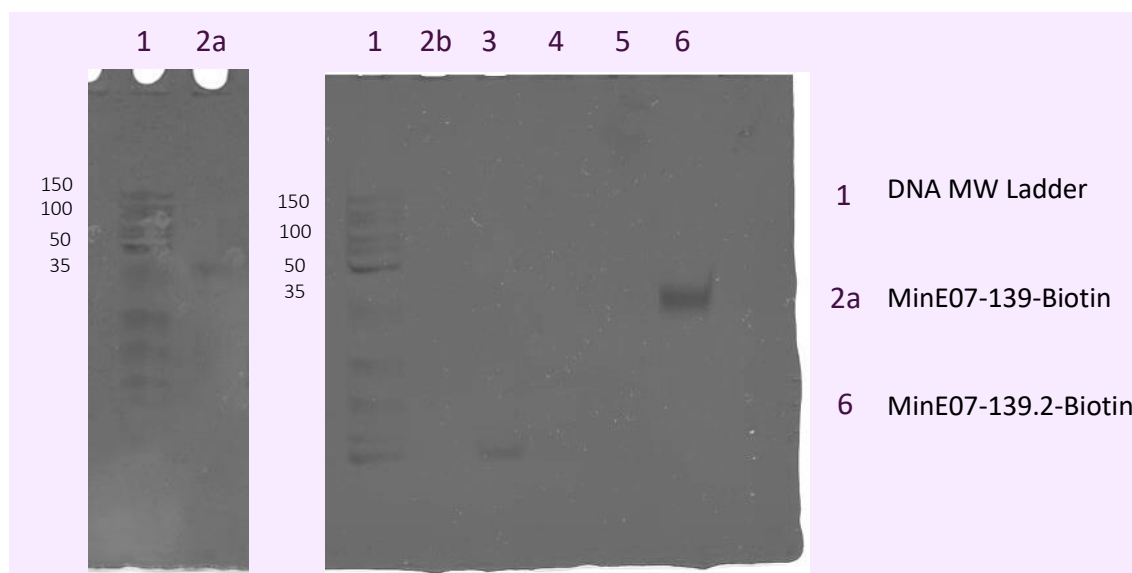


Figure 6.13. 15% Polyacrylamide Gel of MinE07-139-Biotin and MinE07-139.2-Biotin after PAGE purification.

## 6.4.2. HABA ('2-(4′-hydroxyazobenzene) Biotin Binding Assay of PAGE Purified Hit Aptamers.

### 6.4.6.1. Design of HABA Assay and Initial Experiments

To make sure the bands that were cut out during the PAGE purification are the correct ones a HABA assay was designed and run to check they have a biotin on the 5' end of the aptamers. HABA occupies the biotin binding site of streptavidin.[511] In the absence of biotin the HABA will quench the emission from the tropix CPD star. This is because of Fluorescence Resonance Energy Transfer (FRET), (figure 6.14). FRET is a mechanism describing energy transfer between two-sensitive molecules, chromophores.[512] A donor chromophore which is in its electronic excited state will transfer energy to an acceptor chromophore through non-radioactive dipole-dipole coupling.[513] The efficiency of this energy transfer is $10^{-6}$ of the distance between the

acceptor and the donor. This makes FRET extremely sensitive to small changes in distance.[514]

FRET works by having a protein that has two fluorophores attached to it that have overlapping wavelengths. FRET can be using for studying molecular interactions inside living cells. FRET microscopy relies on the ability to capture fluorescent signals from the interactions of labelled molecules in single living or fixed cells. FRET microscopy with cyan fluorescent protein (CFP) fusion protein and a yellow fluorescent protein (YFP) fusion protein FRET pair allows the detection of direct intermolecular integrin interactions in vivo.[515] If FRET occurs the donor channel signal will be quenched and the acceptor channel signal will be sensitised or increased.
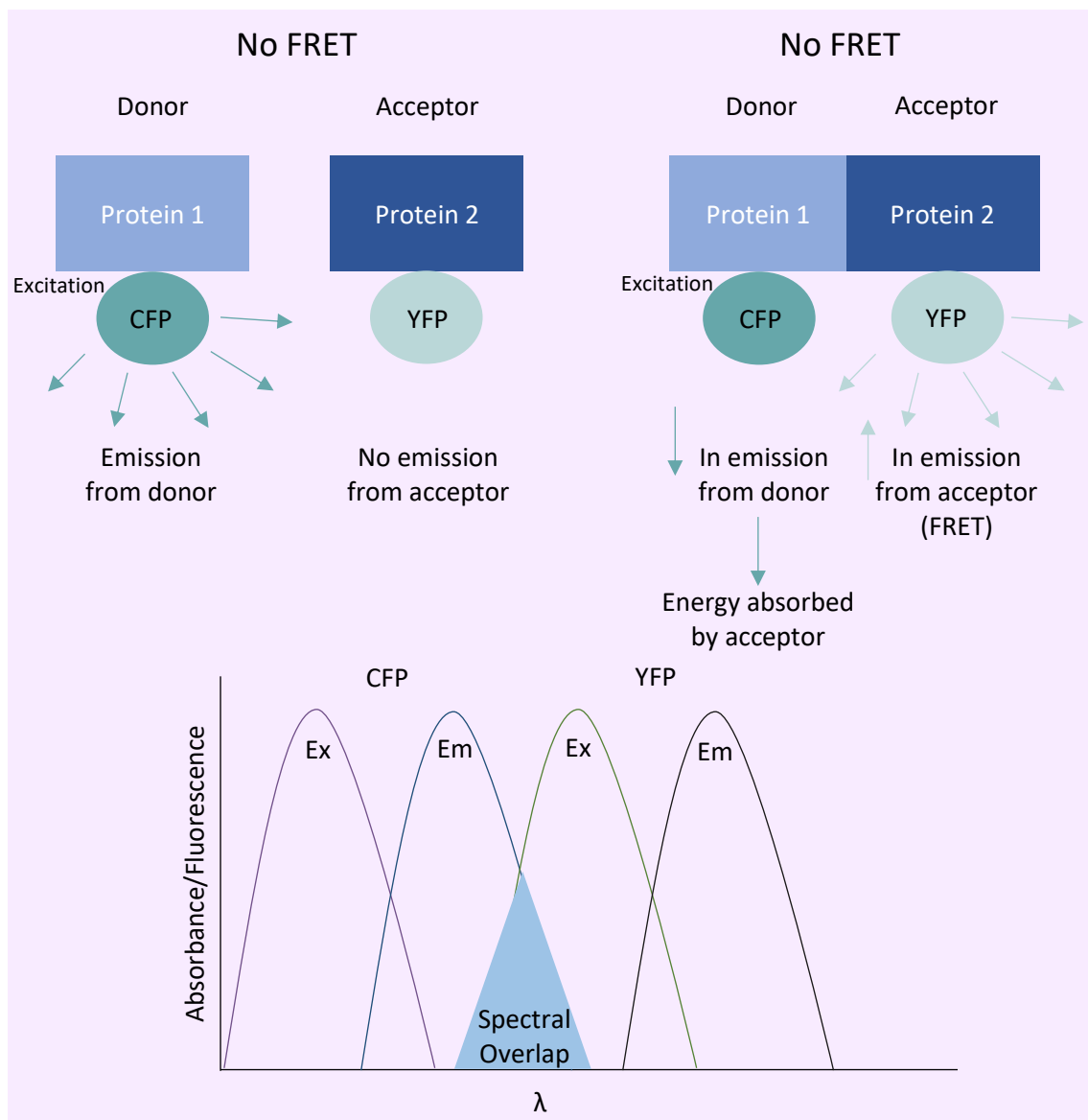


Figure 6.14. Fluorescence Resonance Energy Transfer.

The HABA assay was designed as follows: Streptavidin-Alkaline Phosphatase (Streptavidin-AP) was bound to the plate and the Tropix CDP-Star Ready to use with Sapphire II (Tropix) added to give a chemiluminescence signal. With the addition of HABA the signal should be diminished to background level signal. On the addition of Aptamer-Biotin the HABA should be displaced by the biotin and the chemiluminescence signal should be detectable again. This assay has been done before with Biotective™ Green reagent.[516]   The Biotective Green reagent consists of avidin covalently labelled with the fluorescent dye Alexa Fluor® 488, together with the quencher dye HABA which occupies the biotin binding sites of the avidin. Without any biotin present, the HABA quenches the fluorescence of Alexa Fluor 488 via FRET. When biotin is added, it displaces the HABA from the avidin and the fluorescence of Alexa Fluor 488 dye is restored.[516]
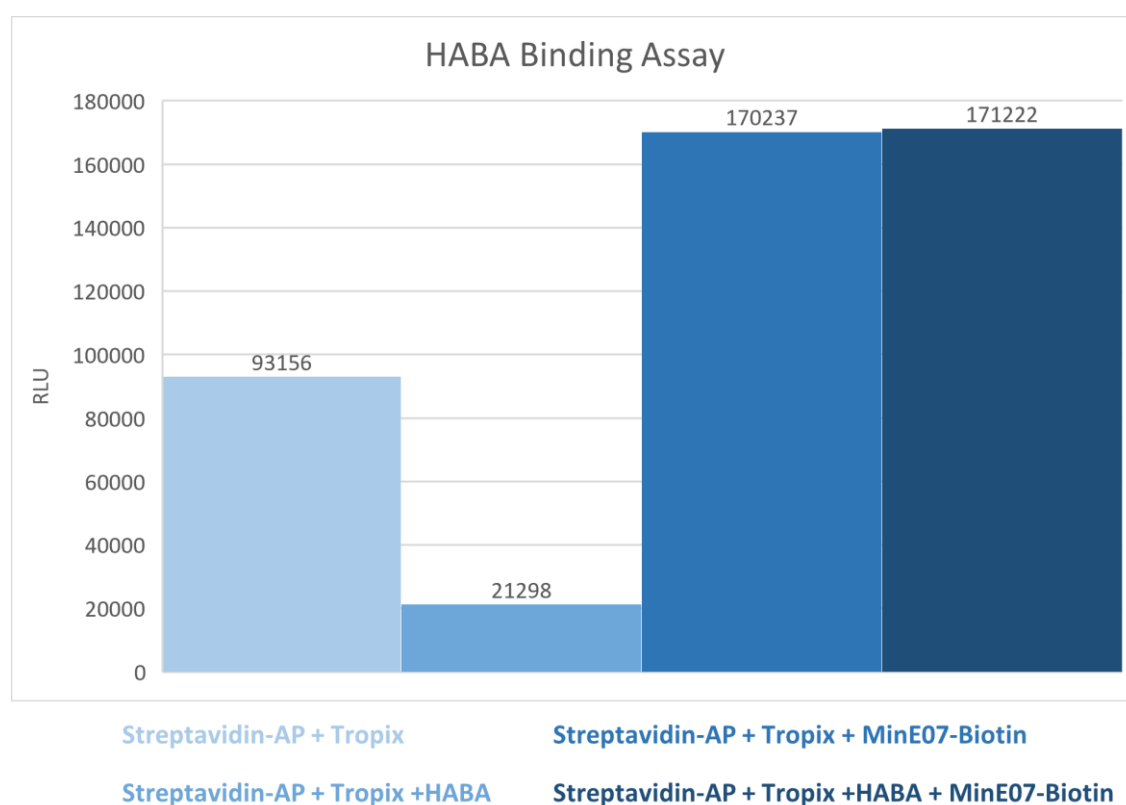


Figure 6.15. Initial HABA Binding Assay proof of concept (n=3).

Figure 6.15 shows the results of the HABA assay run with **MinE07-Biotin-Cent** as a control. The streptavidin-AP incubated together with the Tropix gave signal of 93156 RLU, with the addition

of HABA and the absence of biotin, the signal is diminished to 21298 RLU. With the addition of MinE07-Biotin-Cent the chemiluminescent signal jumps back up to 171222 RLU, which is close to 170237 RLU of the control signal of Streptavidin-AP-Tropix-MinE07-Biotin-Cent with HABA. This initial experiment shows that the biotin displaces the HABA as expected.

### 6.4.6.2. HABA Biotin Binding Assay of Hit and Control Aptamers.

Samples **MinE07M3-Biotin, MinE07M6-Biotin, MinE07U-Ph-Biotin, MinE07U-Vi-Biotin, MinE07U-I-Biotin, MinE07-139-Biotin, MinE07-139.2-Biotin, MinE07-41-Biotin, MinE07-41.2-Biotin, MinE07-56-Biotin, MinE07-139-Biotin** and **MinE07-139.2-Biotin** were then tested in the same way to check that the synthesis and purification was successful. To get a background signal for each modified **MinE07-Biotin** oligomer, Streptavidin-AP was incubated with Tropix and the **MinE07-Biotin** sample. Next Streptavidin-AP was incubated with Tropix and HABA, which gave a diminished signal. The modified **MinE07-Biotin** samples were then added to see if they would displace the HABA to give a higher signal. If the signal was lower than the background signal of Streptavidin-AP incubated with Tropix and the **MinE07-Biotin** sample, it shows that the HABA was not being displaced by a biotin molecule. All modified MinE07 aptamers showed a positive signal, meaning they all have biotin on the 5' end, which could only have occurred if the full-length aptamer was successfully synthesized, due to the nature of the automated synthesis. These samples were carried forward for analysis by protein affinity assay (figure 6.16).

HABA Biotin Binding Assay

| Bar | RLU |
|-----|-----|
| Streptavidin-AP + Tropix | 207668 |
| Streptavidin-AP + Tropix +HABA | 111491 |
| MinE07 | 111613 |
| MinE07-Biotin | 256055 |
| MinE07M3-Biotin | 210966 |
| MinE07M6-Biotin | 213761 |
| MinE07-U-Ph-Biotin | 313423 |
| MinE07-U-Vi-Biotin | 222122 |
| MinE07-U-I-Biotin | 255260 |
| MinE07-41-Biotin | 219291 |
| MinE07-41.2-Biotin | 232777 |
| MinE07-56-Biotin | 313893 |
| MinE07-139-Biotin | 267277 |
| MinE07-139.2-Biotin | 257576 |

**Streptavidin-AP + Tropix**          **MinE07-U-Ph-Biotin**          MinE07-41-Biotin

**Streptavidin-AP + Tropix +HABA**          **MinE07-U-Vi-Biotin**          **MinE07-41.2-Biotin**

**MinE07**          **MinE07-U-I-Biotin**          **MinE07-56-Biotin**

**MinE07-Biotin**          **MinE07-139-Biotin**

**MinE07M3-Biotin**          MinE07-139.2-Biotin

MinE07M6-Biotin

Figure 6.16. HABA Biotin Binding Assay Results for the modified MinE07 Aptamers (n=3).

## 6.4.3. Protein Binding Affinity Assay of MinE07M3-Biotin and MinE07M6-Biotin compared with parent aptamer MinE07-Biotin-Cent

A protein binding affinity assay was run of **MinE07M3-Biotin** and **MinE07M6-Biotin** compared with commercial **MinE07-Biotin-Cent**. The synthesis and purification of **MinE07M3-Biotin** and **MinE07M6-Biotin** can be found in chapter 4.4.3 and 4.4.4. **MinE07M3-Biotin** and **MinE07M6-Biotin** were expected to have a lower affinity for EGFR than the parent **MinE07**. This assay was carried out to check that the protein affinity assay was working as expected. The results of this assay presented in figure 6.17, which shows that what is predicted the results show how

changing **MinE07** by even 3 or 6 nucleotides can make a big difference in the binding affinity for

EGFR.
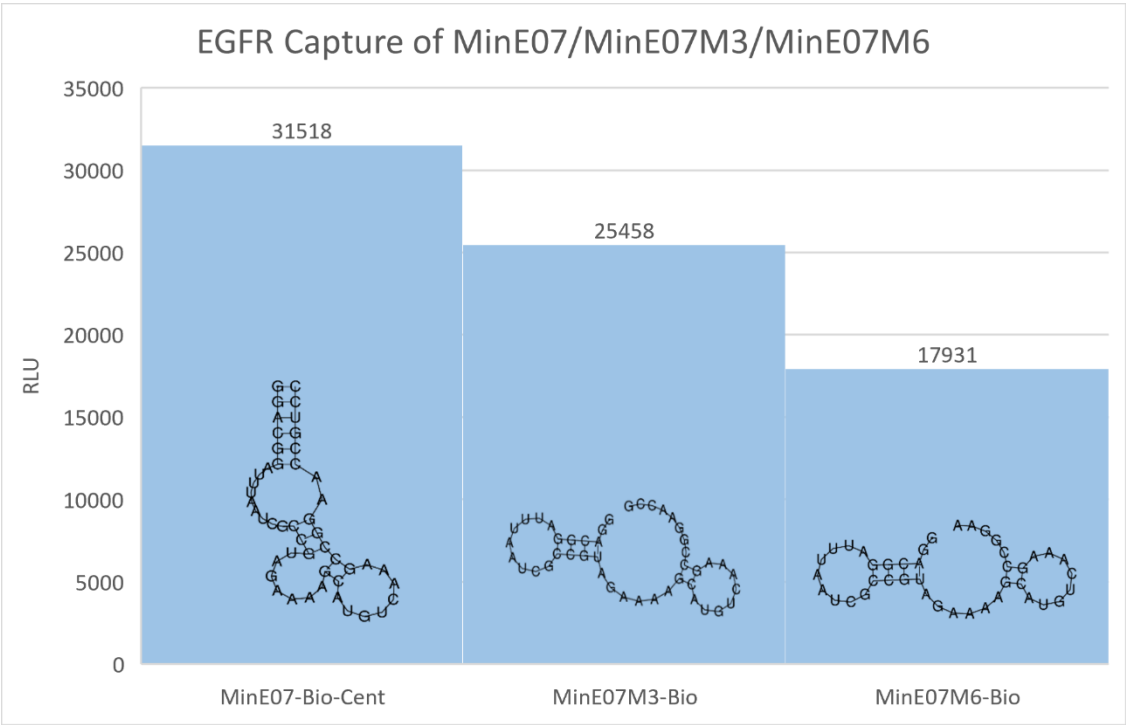


Figure 6.17. Results for Protein Binding Affinity Assay of MinE07M3-Bio and MinE07M6-Bio compared with parent

aptamer MinE07-Bio-Cent (500 nM) (n=4).

## 6.4.4. Protein Binding Affinity Assays of Hit Aptamers

6.4.4.1. Protein Binding Affinity Assays of MinE07-Biotin-Cent

MinE07-Biotin-Cent: 5'-Biotin-fCrGrGrAfUfUfUrArAfUfCrGfCfCrGfUrArGrArArArArGfCrAfUrGfU fCrArArArArGfCfCrGrGrArAfCfCrGfUfCfC-3'.

Three biological repeats were run for each protein binding affinity. Data was analysed using Prism GraphPad and the one site - specific binding model to calculated Bmax and $K_d$ and create the graphs. **MinE07-Biotin-Cent** produced a $K_d$ of 3.696 +/- 0.375 nM (figure 6.18). This is similar to the literature value of aptamer E07 which is 2.4 nM. All other modified **MinE07** aptamers will be compared to this result.
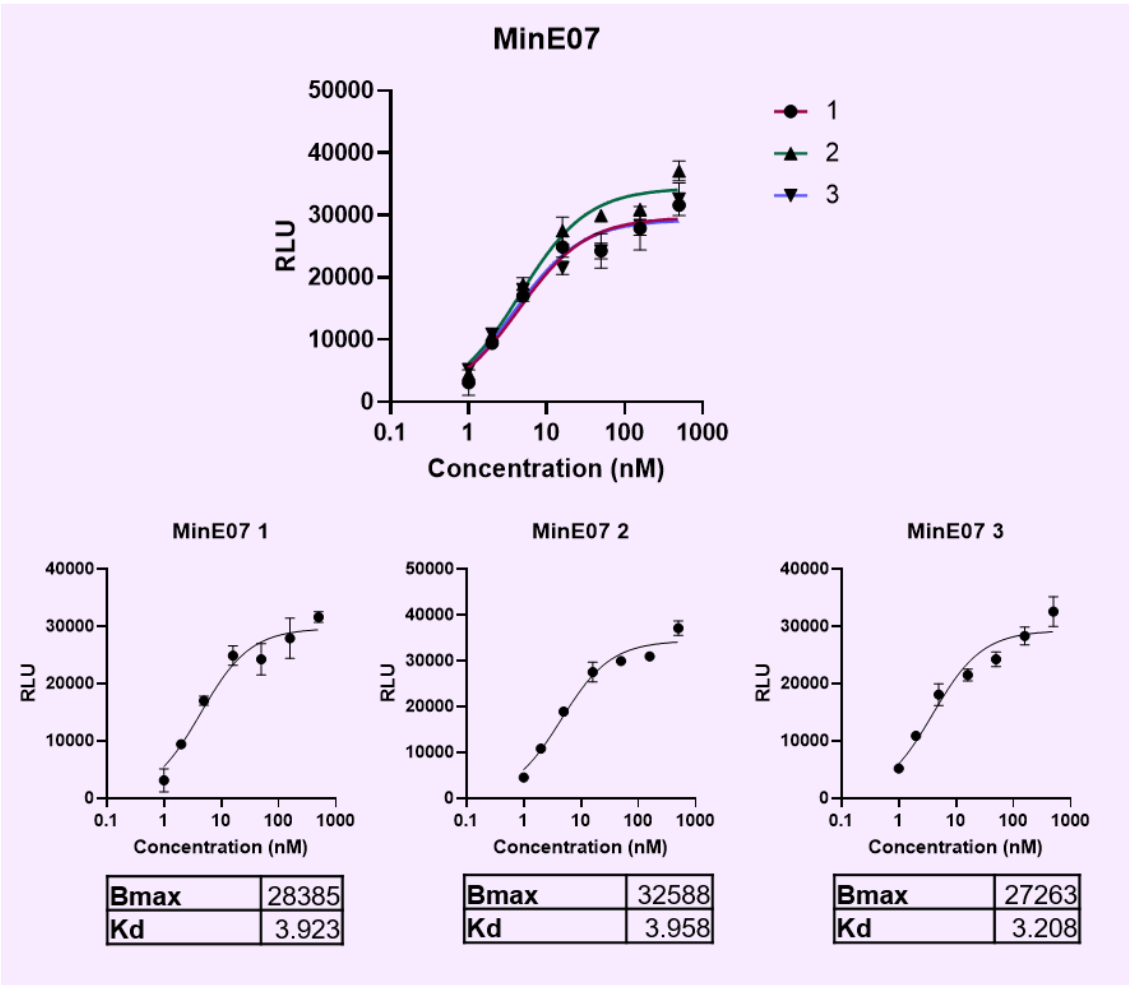


Figure 6.18. Protein Binding Affinity Assay Graphs of MinE07-Biotin-Cent (n=3).

### 6.4.4.2. Protein Binding Affinity Assays of MinE07-U-Ph-Biotin

MinE07-U-Ph-Biotin:5'-Biotin-rGrGrAfCrGrGrAfU-PhfU-PhfU-PhrArAfU-PhfCrGfCfCrGfU-PhrAr

GrArArArArGfCrAfU-PhrGfU-PhfCrArArArArGfCfCrGrGrArAfCfCrGfU-PhfCfC-3'.

**MinE07-U-Ph-Biotin** produced a $K_d$ of 3.590 +/- 0.818 nM (figure 6.19). This result is similar to the parent MinE07 aptamer however the standard deviation is large, potentially meaning the phenyl modification is making it harder to get consistent results. This modification is more sterically demanding that the other two modifications, potentially it would have been expected to have a much lower affinity than MinE07. This could be down to the location of the phenyl modifications; it would be interesting to see if changing their location changed the affinity. A docking study on this compound may give some insight into this.
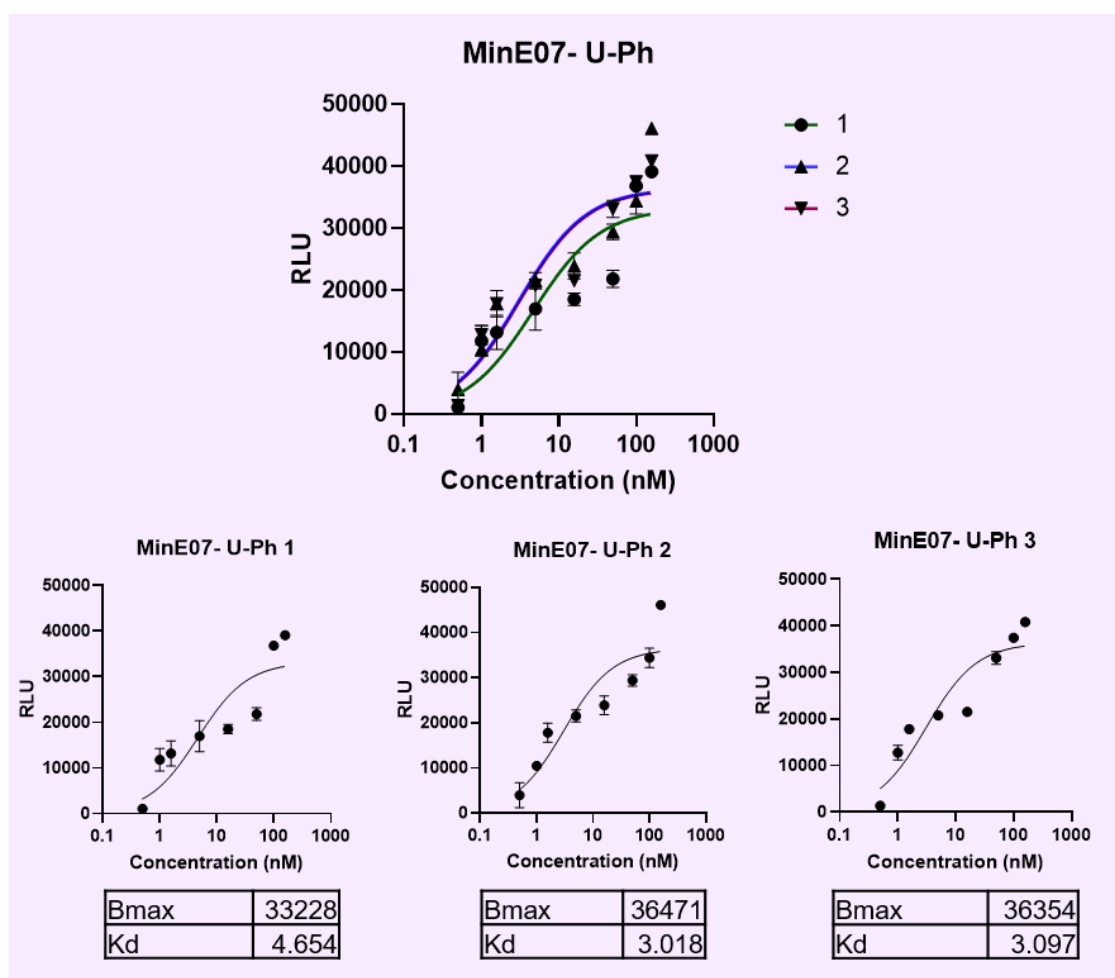


Figure 6.19. Protein Binding Affinity Assay Graphs of MinE07-U-Ph-Biotin (n=3).

6.4.4.3. Protein Binding Affinity Assays of MinE07-U-Vi-Biotin

MinE07-U-Vi-Biotin:5'-Biotin-rGrGrAfCrGrGrAfU-VifU-VifU-VirArAfU-VifCrGfCfCrGfU-VirArGrAr

ArArArGfCrAfU-VirGfU-VifCrArArArGfCfCrGrGrArArAfCfCrGfU-VifCfC-3'

**MinE07-U-Vi-Biotin** produced a $K_d$ of 4.621 +/- 0.482 nM (figure 6.20). This result shows that modifying all the uridines with fU-Vi does not increase the affinity of MinE07s for EGFR. This proves that modifying every uridine will not always improve MinE07 binding to EGFR. The standard deviation is again larger than wanted, potentially meaning that the vinyl modification is making it harder to get consistent results.
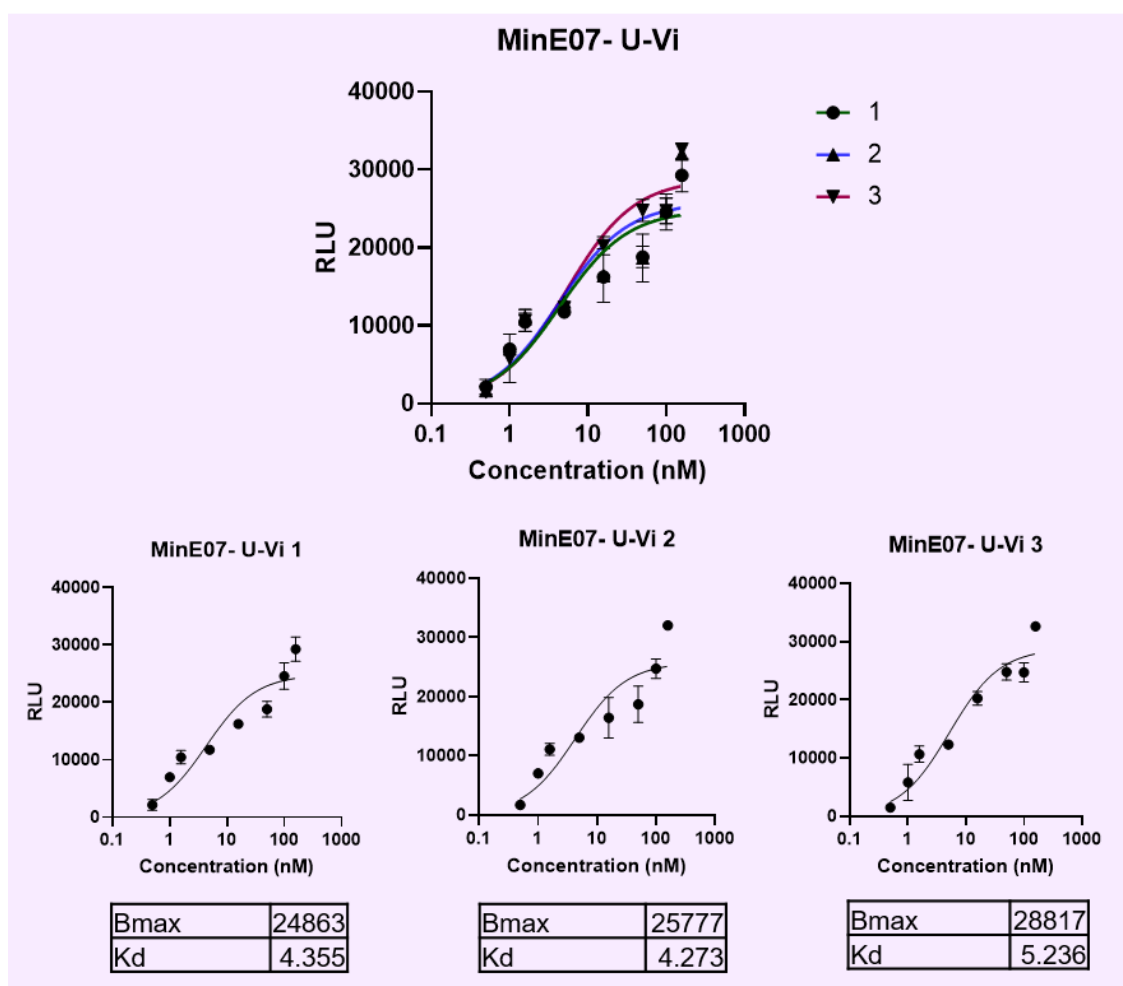


Figure 6.20. Protein Binding Affinity Assay Graphs of MinE07-U-Vi-Biotin (n=3).

### 6.4.4.4. Protein Binding Affinity Assays of MinE07-U-I-Biotin

MinE07-U-I-Biotin:5'-Biotin-rGrGrAfCrGrGrAfU-IfU-IfU-IrArAfU-IfCrGfCfCrGfU-IrArGrArArArArG

fCrAfU-IrGfU-IfCrArArArGfCfCrGrGrArAfCfCrGfU-IfCfC-3'.

**MinE07-U-I-Biotin** produced a $K_d$ of 2.051 +/- 0.256 nM (figure 6.21). Modifying all the uridines with I produced an aptamer with higher affinity for EGFR. The data is more consistent with this modification. This aptamer is the most successful out of the three fully modified aptamers. A significant number of drug candidates in clinical development have halogens in their structures. Insertion of halogen atoms on hit compounds is used to exploit their steric effects, as these bulk atoms can occupy the binding sites of molecular targets. The formation of halogen bonds in ligand-target complexes is recognised as an intermolecular interaction that contributes to the stability of ligand-target complexes.[517]
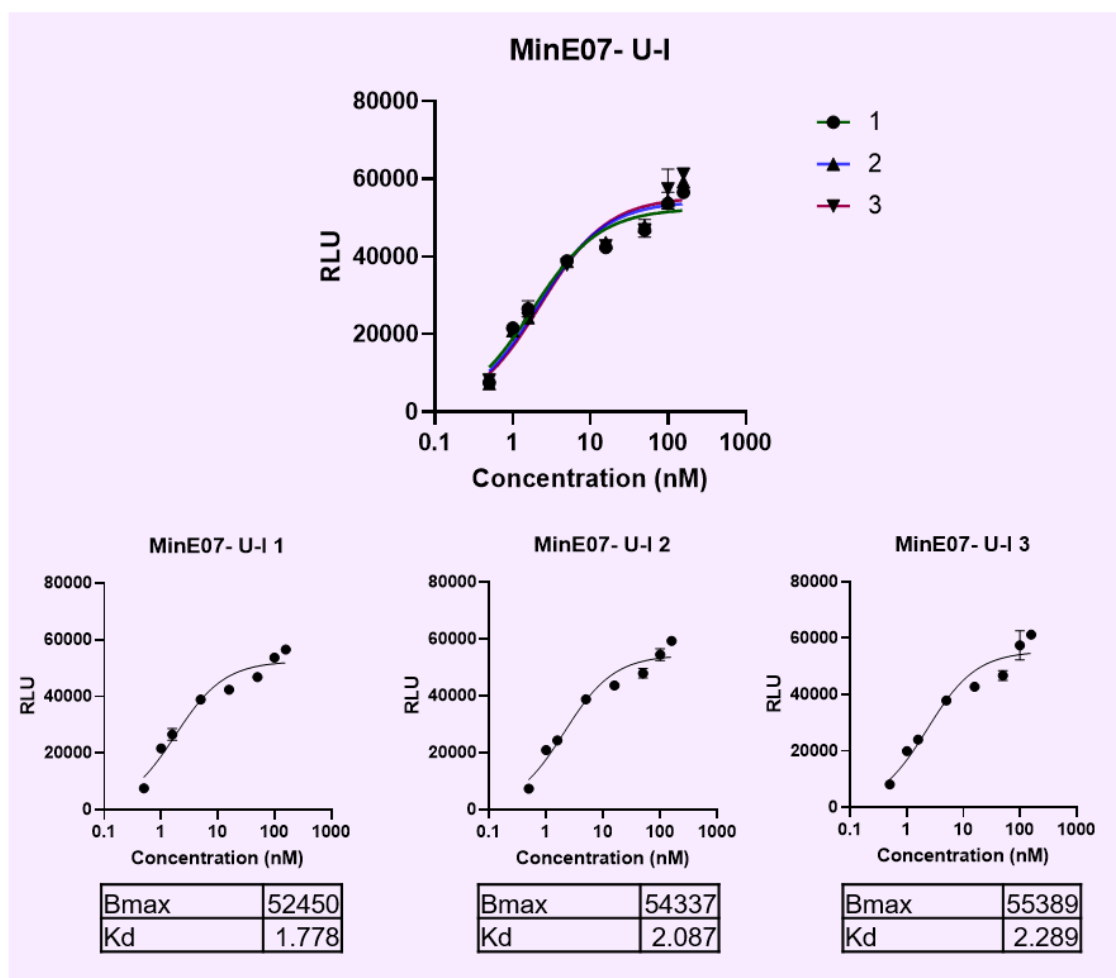
Figure 6.21. Protein Binding Affinity Assay Graphs of MinE07-U-I-Biotin (n=3).

In future work, this fU-I modification could be tried with other aptamers for different targets to see if the results are similar, as these initial results are very promising.

6.4.4.5. Protein Binding Affinity Assays of MinE07-41-Biotin

MinE07-41-Biotin: 5'-Biotin-fCrGrGrAfUfU-PhfU-PhrArAfUfCrGfCfCrGfU-IrArGrArArArArGfCrAf

UrGfUfCrArArArGfCfCrGrGrArAfCfCrGfU-PhfCfC-3'.

**MinE07-41-Biotin** produced a $K_d$ of 1.530 +/- 0.127 nM (figure 6.22). This result shows that the selection and identification process has been successful. The binding affinity of this modified aptamers is greater than for the parent MinE07. The data is also very consistent. This aptamer should go forward for further studies.
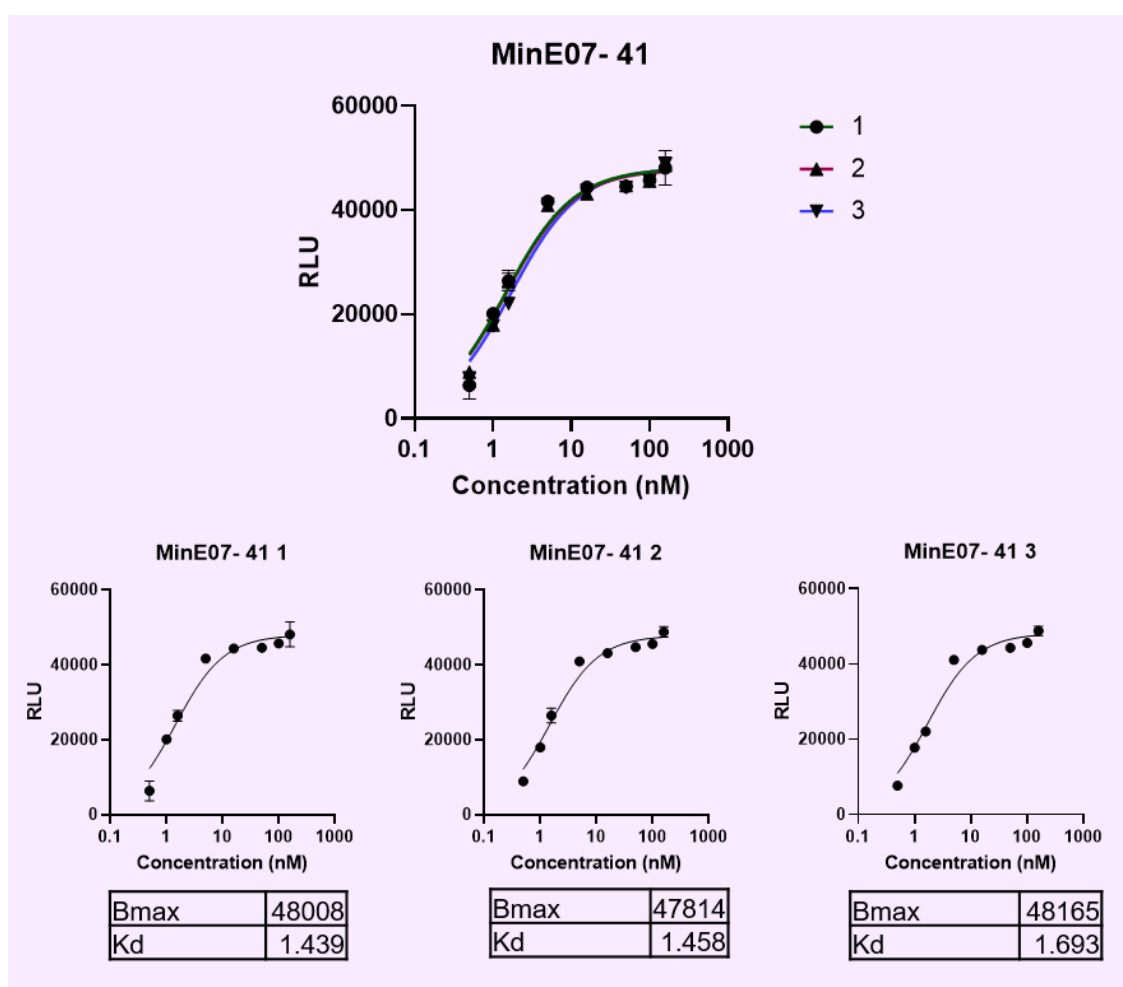


Figure 6.22. Protein Binding Affinity Assay Graphs of MinE07-41-Biotin (n=3).

6.4.4.6. Protein Binding Affinity Assays of MinE07-41.2-Biotin

MinE07-41.2-Biotin: 5'-Biotin-rGrGrAfCrGrGrAfUfU-PhfU-PhrArAfUfCrGfCfCrGfU-IrArGrArArAr
ArGfCrAfUrGfUfCrArArArGfCfCrGrGrArAfCfCrGfU-PhfCfC-3'.

**MinE07-41.2-Biotin** produced a $K_d$ of 3.391 +/- 0.705 nM (figure 6.23). **MinE07-41.2-Biotin**
being the full-length version of **MinE07-41** gives a similar result to the parent MinE07 aptamer.
The standard deviation is high, also suggesting that this aptamer with the extra 3 bases does
not produce consistent results. This would suggest that the 45mer that was identified as a hit
aptamer (**MinE07-41**) was correctly sorted, even though this was shorter, potentially due to
technical error on the synthesiser (premature chain termination).
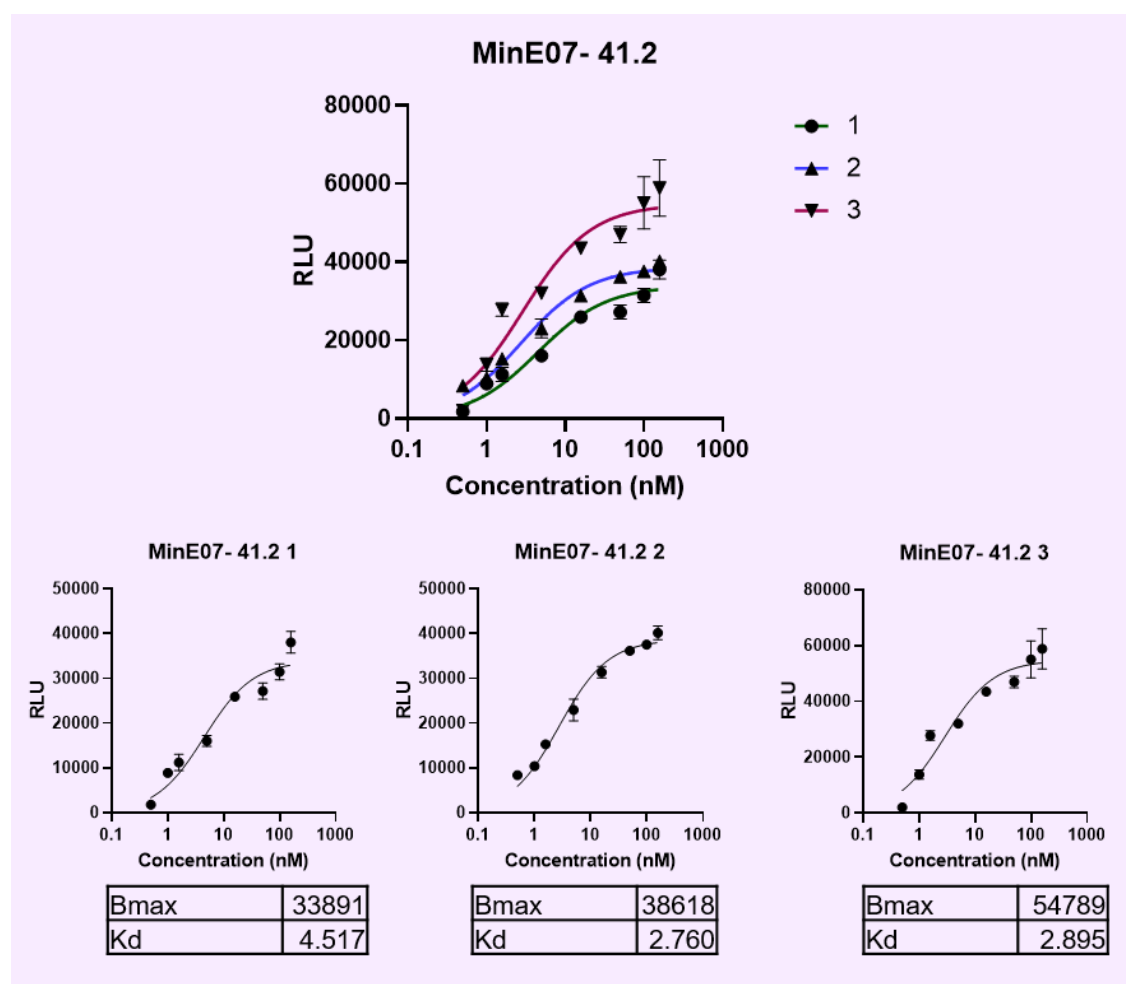


Figure 6.23. Protein Binding Affinity Assay Graphs of MinE07-41.2-Biotin (n=3).

6.4.4.7. Protein Binding Affinity Assays of MinE07-56-Biotin

MinE07-56-Biotin: 5'-Biotin-rGrGrAfCrGrGrAfUfU-PhfU-PhrArAfU-IfCrGfCfCrGfU-PhrArGrArArA
rArGfCrAfUrGfUfCrArArArGfCfCrGrGrArArAfCfCrGfUfCfC-3'

**MinE07-56-Biotin** produced a $K_d$ of 1.468 +/- 0.535 nM (figure 6.24). The results show that the
sorting and identification worked as this modified aptamer does have a higher affinity for EGFR
than the aptamer MinE07 aptamer. The MS/MS results for this aptamer were not completely
conclusive, however the assay results confirm that the proposed **MinE07-56-Biotin** sequence is
a hit aptamer, as it has a higher affinity for EGFR versus MinE07.
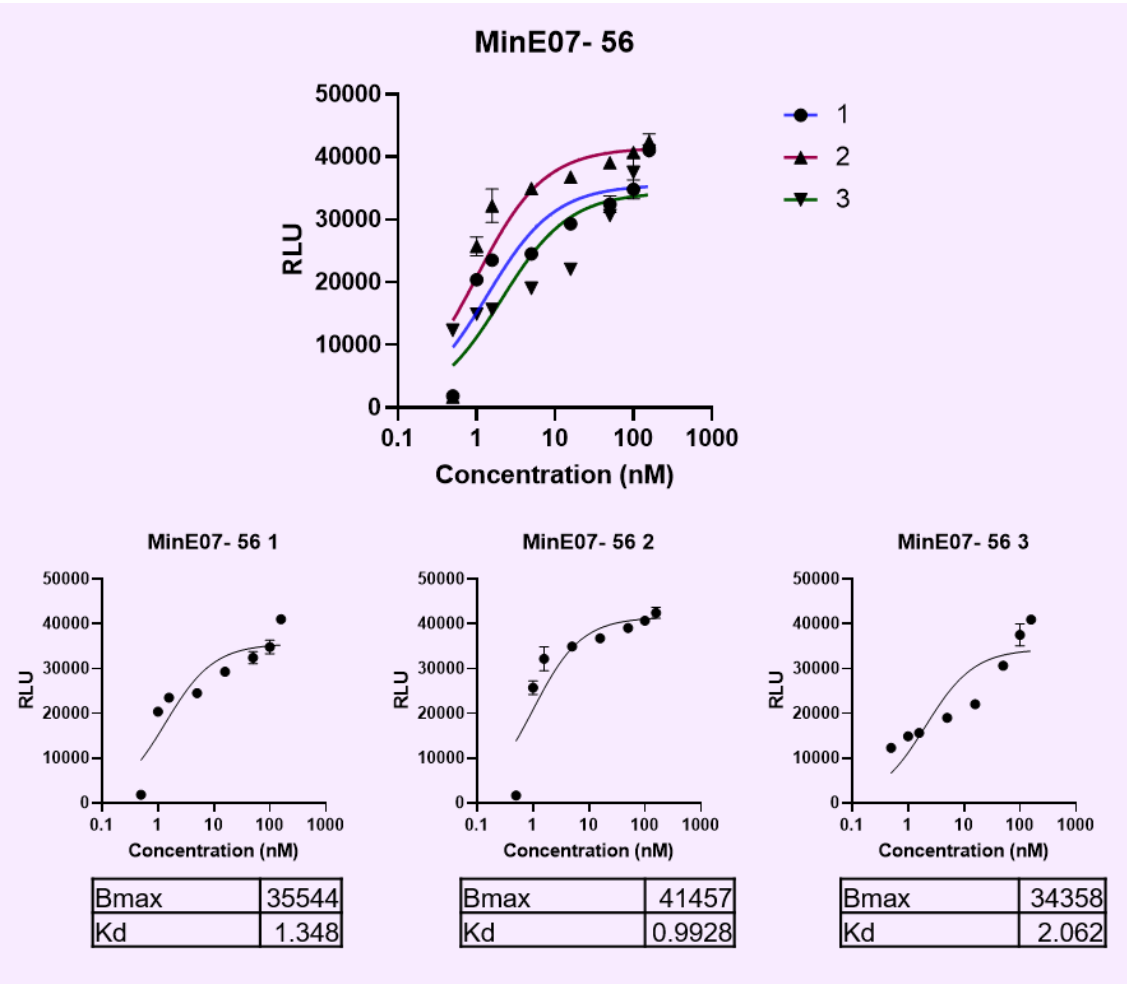


Figure 6.24. Protein Binding Affinity Assay Graphs of MinE07-56-Biotin (n=3).

6.4.4.8. Protein Binding Affinity Assays of MinE07-139-Biotin

MinE07-139-Biotin: 5'-Biotin-rGrGrAfU-VifUfUrArAfUfCrGfCfCrGfUrArGrArArArArGfCrAfU-Phr GfU-VifCrArArArGfCfCrGrGrArAfCfCrGfUfCfC-3'

**MinE07-139-Biotin** produced a $K_d$ of 1.233 +/- 0.099 nM (figure 6.25). The results show that the selection and identification process worked as this modified aptamer, even though it is a 44mer instead of a 48mer because of experimental error (premature chain cleavage during synthesis), has higher affinity for EGFR than the parent aptamer MinE07. The results are very consistent and have a low standard deviation. **MinE07-139** is a very promising EGFR aptamer candidate.
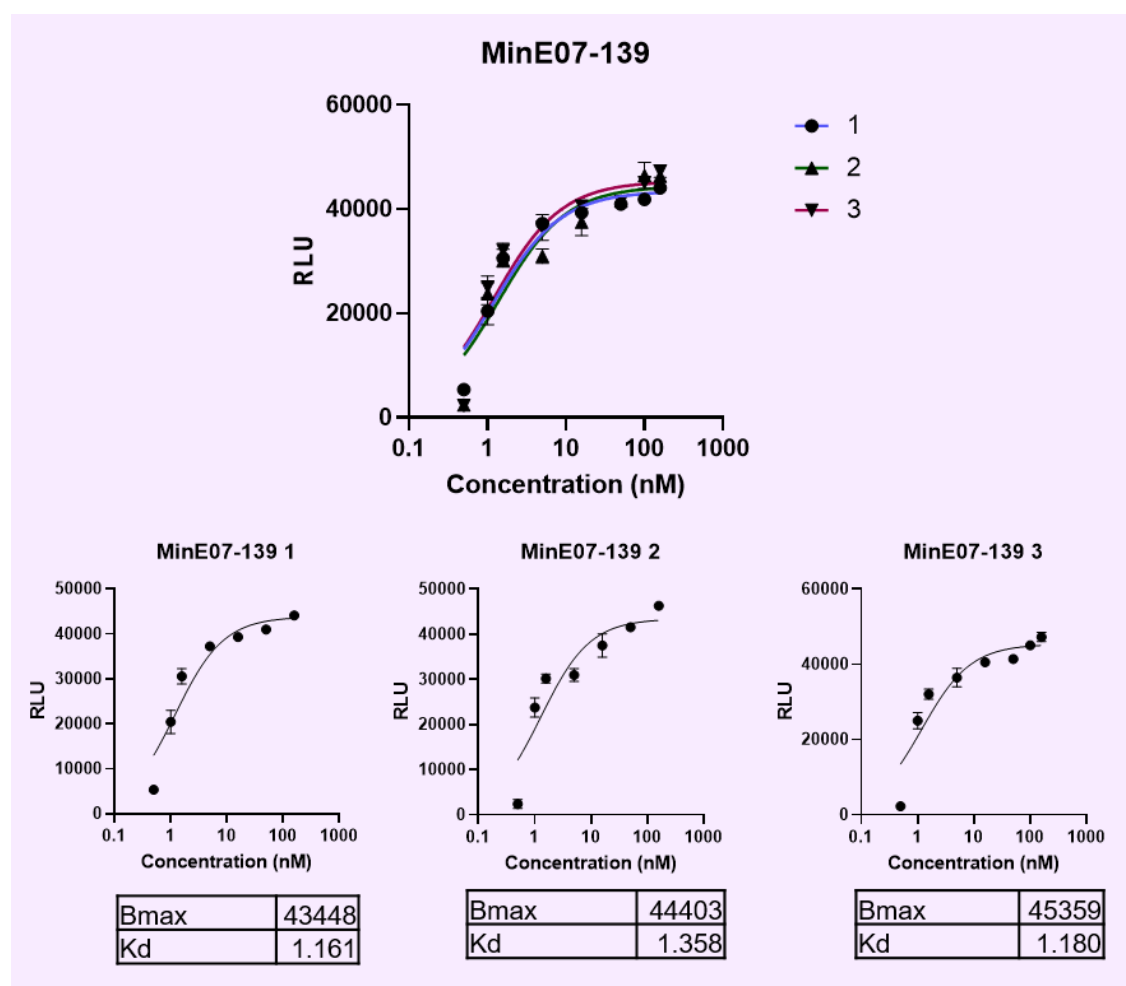


Figure 6.25. Protein Binding Affinity Assay Graphs of MinE07-139-Biotin (n=3).

6.4.4.9. Protein Binding Affinity Assays of MinE07-139.2-Biotin

MinE07-139.2-Biotin: 5'-Biotin-rGrGrAfCrGrGrAfU-VifUfUrArAfUfCrGfCfCrGfUrArGrArArArArGf
CrAfU-PhrGfU-VifCrArArArGfCfCrGrGrArAfCfCrGfUfCfC-3'

**MinE07-139.2-Biotin** produced a $K_d$ of 1.105 +/- 0.447 nM (figure 6.26). The results show that even with experimental error of the one found by sorting and identification (**MinE07-139**), that this method is still valuable because even with minor error the full-length version still produced a modified aptamer with higher affinity for EGFR versus MinE07. The data isn't as consistent as **MinE07-139** however still a potentially great EGFR aptamer candidate.
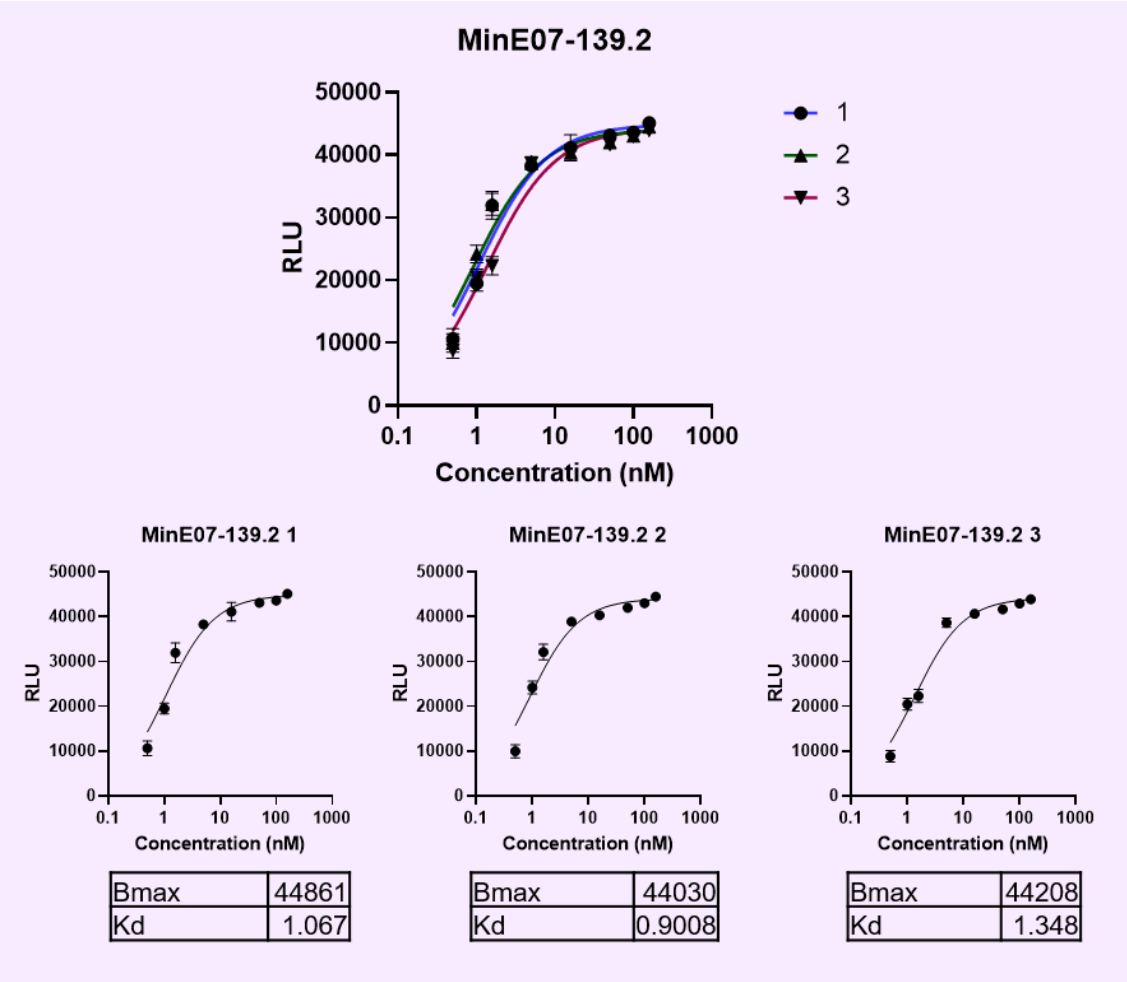


Figure 6.26. Protein Binding Affinity Assay Graphs of MinE07-139.2-Biotin (n=3).

6.4.4.10. Comparision of all $K_d$ values

The results in the table 6.1 below show the $K_d$ values of all the experimentally tested modified aptamers compared with the parent **MinE07** aptamer (**MinE07-Biotin-Cent**).

Table 6.1. $K_d$ values of all experimentally tested modified Aptamers, MinE07-Biotin-Cent, MinE07-U-Ph-Biotin, MinE07-U-Vi-Biotin, MinE07-U-I-Biotin, MinE07-41-Biotin, MinE07-41.2-Biotin, MinE07-56-Biotin, MinE07-139-Biotin and MinE07-139.2-Biotin. Standard deviation shown by +/-. Graphic representation shown in figure 6.27.

| Aptamer | $K_d$ (nM) |
|---------|-----------|
| MinE07-Biotin-Cent | 3.696 +/- 0.375 |
| MinE07-U-Ph-Biotin | 3.590 +/- 0.818 |
| MinE07-U-Vi-Biotin | 4.621 +/- 0.482 |
| MinE07-U-I-Biotin | 2.051 +/- 0.256 |
| MinE07-41-Biotin | 1.530 +/- 0.127 |
| MinE07-41.2-Biotin | 3.391 +/- 0.705 |
| MinE07-56-Biotin | 1.468 +/- 0.535 |
| MinE07-139-Biotin | 1.233 +/- 0.099 |
| MinE07-139.2-Biotin | 1.105 +/- 0.447 |

Comparing all the $K_d$ values of the fully modified and top selected aptamers shows that the selection was essential to the identification of aptamers with a higher affinity for EGFR than the parent MinE07. The position of the modification and the modification itself was important in creating a more successful aptamer. The most promising aptamer being **MinE07-139**, having the most conclusive MS/MS data and protein binding affinity assay results.

**MinE07 Protein Binding Affinity Assay $K_d$ Results**

Legend:
- MinE07-Biotin
- MinE07-U-Ph-Biotin
- MinE07-U-Vi-Biotin
- MinE07-U-I-Biotin
- MinE07-41-Biotin
- MinE07-41.2-Biotin
- MinE07-56-Biotin
- MinE07-139-Biotin
- MinE07-139.2-Biotin
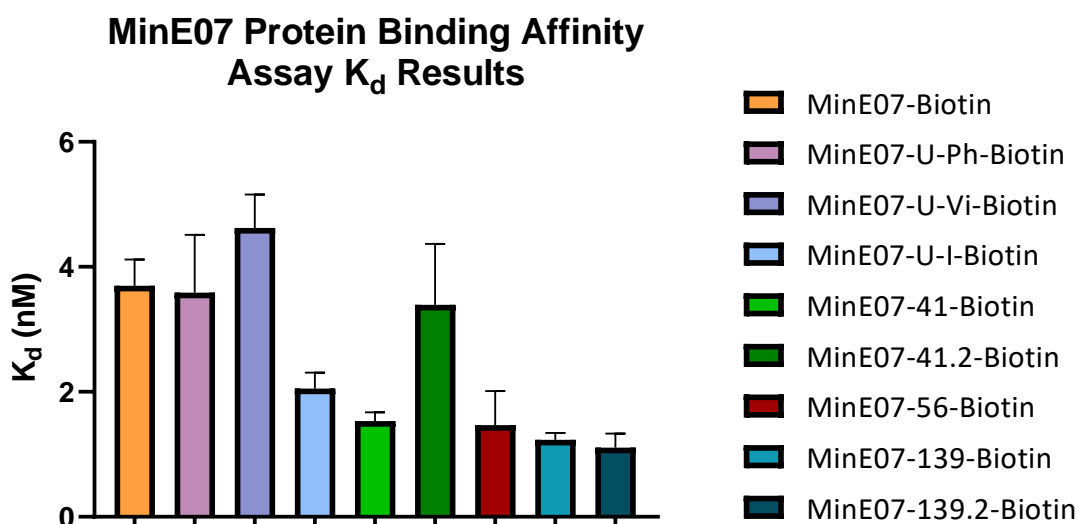
Figure 6.27. Graph of all modified aptamers experimental Kd values (n = 3).

# 6.5. Conclusion

In conclusion, this PhD thesis displays details and success of a novel method for the synthetic selection of enhanced therapeutic aptamers. Starting from the design and synthesis of the solid support to be used for the synthesis of the aptamer library and suitable for being run through the FACS. The modified uridine phosphoramidite molecules fU-Ph, fU-Vi and fU-I were designed and successfully synthesised. The MinE07 aptamer library was successfully synthesised using the split and mix method. Control aptamers were also successfully synthesised**, MinE07-Biotin, MinE07M3-Biotin, MinE07M6-Biotin, MinE07-U-Ph-Biotin, MinE07-U-Vi-Biotin**, and **MinE07-U-I-Biotin.** The MinE07 aptamer library was incubated with EGFR-Fc and Protein A-FITC and successfully sorted using the FACS to give 170 top hit aptamers with the predicted highest affinity to EGFR. Of these hit aptamers, three were successfully identified (**MinE07-41, MinE07-56 and MinE07-139**) using LC-MS/MS and data analysis with RoboOligo software. **MinE07-41-Biotin, MinE07-56-Biotin** and **MinE07-139-Biotin** were re-synthesised for validation studies and to determine their $K_d$ values for EGFR. Protein affinity binding assays were run on **MinE07-Biotin-Cent, MinE07-U-Ph-Biotin, MinE07-U-Vi-Biotin, MinE07-U-I-Biotin, MinE07-41-Biotin,**

**MinE07-41.2-Biotin, MinE07-56-Biotin, MinE07-139-Biotin** and **MinE07-139.2-Biotin**. $K_d$ values were generated providing evidence that this novel method for the synthetic selection of enhanced therapeutic aptamers was successful. Five novel aptamers have been identified as having higher affinity for EGFR (**MinE07-U-I-Biotin, MinE07-41-Biotin, MinE07-56-Biotin, MinE07-139-Biotin** and **MinE07-139.2-Biotin**) versus the parental MinE07 aptamer.

## 6.6. Potential Research Outputs

This project has the potential to impact: (i) drug discovery processes with this new method of synthesising and screening large drug candidate libraries rapidly and (ii) EGFR targeted therapeutics by providing new aptamer drug candidates

Although thousands of excellent aptamers targeting a broad range of targets have been generated by SELEX experiments since 1990, outstanding aptamers that can be reliably used in biomedical and analytical applications to solve real-world problems are still limited. This is because there are still a gaps in their performance including their specificity and affinity.[518] To overcome these hurdles, researchers have been suggesting that there needs to be an innovative method to create aptamers with large library sizes and more complex structures.[519] This research produces an innovative method to increase modifications and complexity into an aptamer library, to potentially increase affinity and specificity. Current research shows that DNA and RNA aptamers have a median $K_d$ value of 32.8 nM and 19.7 nM respectively.[519] This novel method of modifying nucleic acids that has been displayed in this thesis has been producing aptamers in the range of 1-5 nM $K_d$. Consequently, this research could drive the progress of some aptamers that have already been selected, whose affinity for their target was failing their progress into clinical trials. A typical SELEX process takes 2-3 months for selection and sequence identification.[60] This research project did the aptamer selection in 1 day followed by 2-4 weeks

of sequence analysis. This novel method could improve the success of aptamer discovery processes and general drug discovery process, making them more cost and time effective.

The aptamer market is forever growing, as research and industry see the advantages of them. The global aptamer market size is estimated to value USD 242.3 million in 2022 and grow to USD 523.5 million in 2027.[520] The primary factors driving the growth are technological advancements and increasing investments from the pharmaceutical and biotechnology industries.[521] The low cost and high efficacy in binding to large molecules compared to antibodies as well as the patent expiration of SELEX has boosted this aptamer market.[522] In 2019, the Aptamer Group announced that AstraZeneca Company has agreed to collaborate to expand the aptamer technology application potential. AstraZeneca will adapt the aptamer technology to improve their drug delivery vehicles.[523] This novel project displayed in this thesis is really relevant to the science advancements and future of aptamers.

In the age of personalised cancer medicine, targeted therapies have become an essential part of cancer treatment in combination with chemotherapy and radiotherapy. Many molecular targeted drug are being investigated and are in the drug development process.[524] Aptamers as cancer therapeutics have evolved rapidly in the past 10 years which has been shown by the increase in peer-review publications and patents involving aptamer anti-cancer drugs. A variety of molecules involving tumour progression and metastasis have been targeted by aptamers, in various locations such as cancer cells, tumour associated vessels and tumour stroma. Pre-clinical testing shows aptamers have promising efficacies as well have having good toxicity levels. There are currently two anti-cancer aptamers in clinical trials, as safe, high affinity therapeutics, aptamers present an excellent opportunity to make personalised medicine a reality for more cancer patients in the future.[525]

The efficacy of first, second and third generation TKIs depends on the mutation that EGFR has. Around 70% of all EGFR-mutated tumours respond clinically to TKIs, seeing considerable tumour

size shrinkage in the present of these drugs. The other 30% are resistant to TKI drugs and so do not respond to treatment.[526] Even with these great results the TKIs are causing resistance and have substantial toxicity, causing side effects such as diarrhoea, skin rashes, paronychia and stomatitis. This has required doses to be reduced by 40% in these patients.[527] Aptamers have been suggested to help battle resistance [528] as well as already being successful in previous clinical trials (Pegaptanib)[529] as they give minimal side effects because of its nature of high specificity and affinity for their targets,[518] showing that aptamers have the potential to be a successful ligand for EGFR.

## 6.7. Future Work

In future work, this novel selection method should undergo further research and development as there are many possibilities for uses in therapeutic discovery, drug delivery vehicles and diagnostics. Some extra negative selection steps should be added into the method to ensure that the new modifications on the aptamers are not making it less selective. Experimentally this could be done by getting some other proteins that are highly expressed on cancer cells and run binding studies with them and the modified MinE07 aptamers. If moving onto *in vitro* work then different cell types could be compared, some that highly express EGFR and some that do not, and hopefully this will show the aptamers are selective for EGFR.

This novel selection method should be developed to run smoother and simpler. This could be done by picking other potentially successful aptamers and improving the method while also selecting out better aptamer candidates. This method could also be tried in reverse by first selecting a protein or cellular target and synthesising a modified aptamer library to do a selection with. A larger library with increasing diversity could be synthesised introducing novel modifications to the nucleotides, not just to the uridines.

The novel aptamers that have been discovered, **MinE07-U-I, MinE07-41, MinE07-56, MinE07-139** and **MinE07-139.2-Biotin** could undergo further biological and biophysical testing. Such as NMR spectroscopy binding studies (Saturation Transfer Difference NMR, Carr-Purcell-Meiboom-Gill NMR and Water–ligand observed gradient spectroscopy NMR) which will provide useful structural data for the binding of the aptamers to EGFR, helping to work out the effects of the different modifications and identify ligand binding sites. Molecular dynamics and protein docking studies would also help identify the ligand binding site and produce data that shows why certain modifications in certain places are essential to the increased binding affinity. Cell targeting and gene transfer studies could also be undertaken with the novel aptamers on different types of cancerous cells such as breast, lung and glioblastoma cell lines.
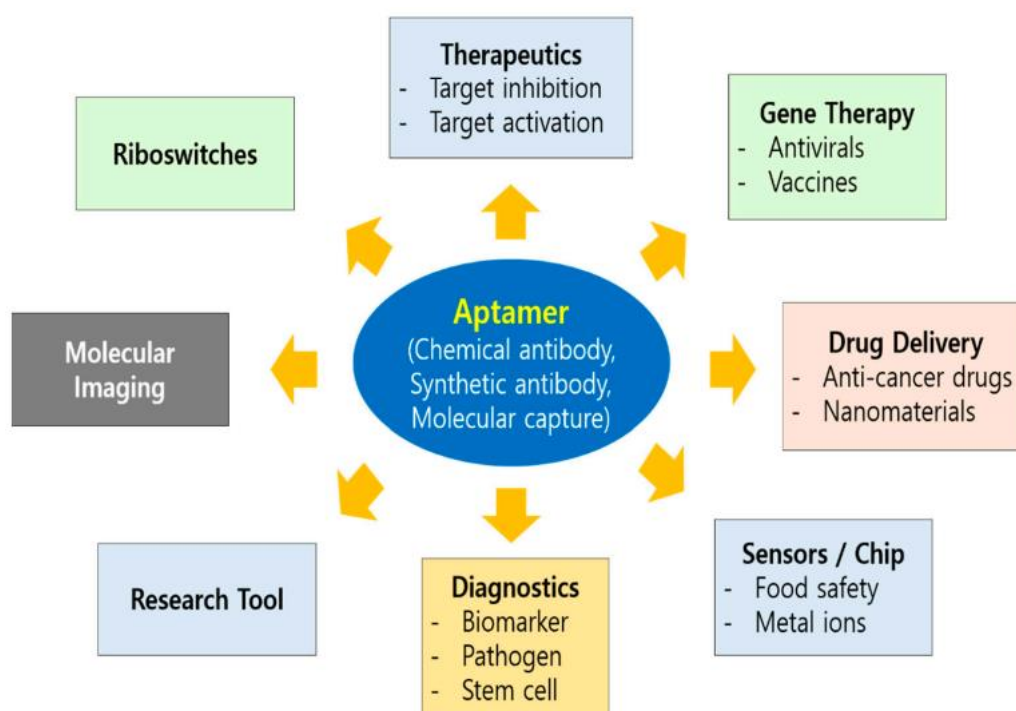


Figure 6.28. Diverse areas of aptamer applications

These highly targeting novel aptamers could be used as drug delivery vehicles. A known drug/therapeutic agent could be conjugated to the aptamer and delivered to EGFR in a highly

targeted approach. Using this aptamer as a drug delivery vehicle would be an interesting project to undertake.

# 6.8. References

1       S. Sigismund, D. Avanzato and L. Lanzetti, *Mol. Oncol.*, 2018, **12**, 3–20.

2       G. Bethune, D. Bethune, N. Ridgway and Z. Xu, *J. Thorac. Dis.*, 2010, **2**, 48–51.

3       J. Wykosky, T. Fenton, F. Furnari and W. K. Cavenee, *Chin. J. Cancer*, 2011, **30**, 5–12.

4       M. R. Dunn, R. M. Jimenez and J. C. Chaput, *Nat. Rev. Chem.*, 2017, **1**, 0076.

5       F. Odeh, H. Nsairat, W. Alshaer, M. A. Ismail, E. Esawi, B. Qaqish, A. A. Bawab and S. I. Ismail, *Molecules*, 2019, **25**, 1–51.

6       N. Li, H. H. Nguyen, M. Byrom and A. D. Ellington, *PLoS One*, 2011, **6**, 1–10.

7       A. D. Keefe, S. Pai and A. Ellington, *Nat. Rev. Drug Discov.*, 2010, **9**, 537–550.

8       J. H. P. Chan, S. Lim and W. S. F. Wong, *Clin. Exp. Pharmacol. Physiol.*, 2006, **33**, 533–540.

9       X. Zou, J. Wu, J. Gu, L. Shen and L. Mao, *Front. Microbiol.*, 2019, **10**, 1462.

10      J. C. Cox, A. Hayhurst, J. Hesselberth, T. S. Bayer, G. Georgiou and A. D. Ellington, *Nucleic Acids Res.*, 2002, **30**, 1–14.

11      M. Sprinzl, M. Milovnikova and C. S. Voertler, *Handb. Exp. Pharmacol.*, 2006, **173**, 327–340.

12      C. L. A. Hamula, H. Zhang, L. L. Guan, X. F. Li and X. C. Le, *Anal. Chem.*, 2008, **80**, 7812–7819.

13      K. L. Hong and L. J. Sooter, *Biomed Res. Int.*, 2015, **2015**, 31.

14      T. Takahashi, K. Tada and H. Mihara, *Mol. Biosyst.*, 2009, **5**, 986–991.

15      A. Ruscito and M. C. DeRosa, *Front. Chem.*, 2016, **4**, 1–14.

16      C. Mannironi, C. Scerch, P. Fruscoloni and G. P. Tocchini-Valentini, *Rna*, 2000, **6**, 520–527.

17      S. Catuogno and C. L. Esposito, *Biomedicines*, 2017, **5**, 49.

18      B. Carlson, *Biotechnol. Healthc.*, 2007, **4**, 31–6.

19      Q. Zhu, G. Liu and M. Kai, *Molecules*, 2015, **20**, 20979–20997.

20      E. W. Orava, N. Cicmil and J. Gariépy, *Biochim. Biophys. Acta - Biomembr.*, 2010, **1798**, 2190–2200.

21    S. Reverdatto, D. Burz and A. Shekhtman, *Curr. Top. Med. Chem.*, 2015, **15**, 1082–1101.

22    B. A. Cohen, P. Colas and R. Brent, *Proc. Natl. Acad. Sci. U. S. A.*, 1998, **95**, 14272–14277.

23    I. C. Baines and P. Colas, *Drug Discov. Today*, 2006, **11**, 334–341.

24    Y. Alhamhoom, H. M. As Sobeai, S. Alsanea and A. Alhoshani, *Int. J. Oncol.*, 2022, **60**, 1–15.

25    Y. Ding and C. E. Lawrence, *Nucleic Acids Res.*, 2003, **31**, 7280–7301.

26    C. S. Lim and C. M. Brown, *Front. Microbiol.*, 2018, **8**, 2582.

27    T. Sakamoto, E. Ennifar and Y. Nakamura, *Biochimie*, 2018, **145**, 91–97.

28    T. Mairal, V. Cengiz Özalp, P. Lozano Sánchez, M. Mir, I. Katakis and C. K. O'Sullivan, *Anal. Bioanal. Chem.*, 2008, **390**, 989–1007.

29    T. Hermann and D. J. Patel, *Science (80-. ).*, 2000, **287**, 820–825.

30    S. Nagatoishi, N. Isono, K. Tsumoto and N. Sugimoto, *ChemBioChem*, 2011, **12**, 1822–1826.

31    J. A. Cowan, T. Ohyama, D. Wang and K. Natarajan, *Nucleic Acids Res.*, 2000, **28**, 2935–2942.

32    R. E. Anna Aviñóa, Stefania Mazzinib, Ruben Ferreiraa, Raimundo Gargalloc, Victor E. Marquezd, *Bioorg Med Chem*, 2014, **23**, 1–7.

33    B. Pagano, L. Martino, A. Randazzo and C. Giancola, *Biophys. J.*, 2008, **94**, 562–569.

34    P. T. Lang, S. R. Brozell, S. Mukherjee, E. F. Pettersen, E. C. Meng, V. Thomas, R. C. Rizzo, D. A. Case, T. L. James and I. D. Kuntz, *Rna*, 2009, **15**, 1219–1230.

35    S. B. Long, M. B. Long, R. R. White and B. A. Sullenger, *Rna*, 2008, **14**, 2504–2512.

36    W. T. Horn, M. A. Convery, N. J. Stonehouse, C. J. Adams, L. Liljas, S. E. V. Phillips and P. G. Stockley, *Rna*, 2004, **10**, 1776–1782.

37    R. Amano, K. Takada, Y. Tanaka, Y. Nakamura, G. Kawai, T. Kozu and T. Sakamoto, *Biochemistry*, 2016, **55**, 6221–6229.

38    A. Rangnekar, A. M. Zhang, S. S. Li, K. M. Bompiani, M. N. Hansen, K. V. Gothelf, B. A. Sullenger and T. H. LaBean, *Nanomedicine Nanotechnology, Biol. Med.*, 2012, **8**, 673–681.

39    P. H. Lin, R. H. Chen, C. H. Lee, Y. Chang, C. S. Chen and W. Y. Chen, *Colloids Surfaces B Biointerfaces*, 2011,

**88**, 552–558.

40      L. Xu, L. Cai, X. Chen, X. Jiang, H. Chong, B. Zheng, K. Wang, J. He, W. Chen, T. Zhang, M. Cheng, Y. He and K. Liu, *Antimicrob. Agents Chemother.*, 2013, **57**, 4963–4970.

41      J. C. Rohloff, A. D. Gelinas, T. C. Jarvis, U. A. Ochsner, D. J. Schneider, L. Gold and N. Janjic, *Mol. Ther. - Nucleic Acids*, 2014, **3**, e201.

42      X. Ren, A. D. Gelinas, I. Von Carlowitz, N. Janjic and A. M. Pyle, *Nat. Commun.*, 2017, **8**, 810.

43      S. Cai, J. Yan, H. Xiong, Y. Liu, D. Peng and Z. Liu, *Analyst*, 2018, **143**, 5317–5338.

44      J. Bernard Da Costa and T. Dieckmann, *Mol. Biosyst.*, 2011, **7**, 2156–2163.

45      C. Zhang, L. Wang, Z. Tu, X. Sun, Q. He, Z. Lei, C. Xu, Y. Liu, X. Zhang, J. Yang, X. Liu and Y. Xu, *Biosens. Bioelectron.*, 2014, **55**, 216–219.

46      Y. Nomura, S. Sugiyama, T. Sakamoto, S. Miyakawa, H. Adachi, K. Takano, S. Murakami, T. Inoue, Y. Mori, Y. Nakamura and H. Matsumura, *Nucleic Acids Res.*, 2010, **38**, 7822–7829.

47      Aptamer Binding Site, https://www.idtdna.com/pages/education/decoded/article/planning-to-work-with-aptamers.

48      A. S. R. Potty, K. Kourentzi, H. Fang, P. Schuck and R. C. Willson, *Int. J. Biol. Macromol.*, 2011, **48**, 392–397.

49      M. B. Murphy, S. T. Fuller, P. M. Richardson and S. A. Doyle, *Nucleic Acids Res.*, 2003, **31**, 1–8.

50      R. R. White, B. A. Sullenger, P. Christopher, J. C. Invest, R. R. White, B. A. Sullenger and C. P. Rusconi, 2000, **106**, 929–934.

51      G. C. Roberts, E. A. Dennis, D. H. Meadows, J. S. Cohen and O. Jardetzky, *Proc. Natl. Acad. Sci. U. S. A.*, 1969, **62**, 1151–8.

52      H. Rastogi and D. A. Usher, *Nucleic Acids Res.*, 1995, **23**, 4872–4877.

53      K. Germer, M. Leonard and X. Zhang, *Int. J. Biochem. Mol. Biol.*, 2013, **4**, 27–40.

54      E. Fattal, H. Hillaireau and S. I. Ismail, *Adv. Drug Deliv. Rev.*, 2018, **134**, 1–2.

55      Y. Takafuji, J. I. Jo and Y. Tabata, *J. Biomater. Sci. Polym. Ed.*, 2011, **22**, 1179–1195.

56      A. V. Lakhin, V. Z. Tarantul and L. V. Gening, *Acta Naturae*, 2013, **5**, 34–43.

57    S. Shigdar, J. Luczo, M. Q. Wei, R. Bell, A. Danks, K. Liu and W. Duan, *Open Conf. Proc. J.*, 2010, **1**, 118–124.

58    F. M. Veronese, R. Largajolli and E. Boccu, *Appl. Biochem. Biotechnol.*, 1985, **11**, 141–152.

59    C. A. Trujillo, A. A. Nery, J. M. Alves, A. H. Martins and H. Ulrich, *Clin. Ophthalmol.*, 2007, **1**, 393–402.

60    R. Stoltenburg, C. Reinemann and B. Strehlitz, *Biomol. Eng.*, 2007, **24**, 381–403.

61    S. Diafa and M. Hollenstein, *Molecules*, 2015, **20**, 16643–16671.

62    J. Szostak and M. Sassanfar, *Nature*, 1993, **364**, 550.

63    R. . Mellor, J. Ronnenberg, C. W.H. and S. Diekmann, *Nature*, 1992, **355**, 717–719.

64    J. Liu and G. D. Stormo, *Nucleic Acids Res.*, 2005, **33**, 1–6.

65    R. Stoltenburg, C. Reinemann and B. Strehlitz, *Anal. Bioanal. Chem.*, 2005, **383**, 83–91.

66    M. Bianchini, M. Radrizzani, M. G. Brocardo, G. B. Reyes, C. Gonzalez Solveyra and T. A. Santa-Coloma, *J. Immunol. Methods*, 2001, **252**, 191–197.

67    S. D. Mendonsa and M. T. Bowser, *Anal. Chem.*, 2004, **76**, 5387–5392.

68    R. Y. L. Tsai and R. R. Reed, *Mol. Cell. Biol.*, 1998, **18**, 6447–6456.

69    K. A. Davis, B. Abrams, Y. Lin and S. D. Jayasena, *Nucleic Acids Res.*, 1996, **24**, 702–706.

70    M. Djordjevic, *Biomol. Eng.*, 2007, **24**, 179–189.

71    Aptamers, drug delievery, http://www.scienceinafrica.com/biotechnology/industrial/nuts-and-bolts-genius-behind-aptamers-future-drug-delivery.

72    N. N. Quang, G. Perret and F. Ducongé, *Pharmaceuticals*, 2016, **9**, 1–15.

73    S. S. Jamuar, A. M. D'Gama and C. A. Walsh, *Genomics, Circuits, Pathways Clin. Neuropsychiatry*, 2016, **12**, 1504–1514.

74    K. V. Voelkerding, S. A. Dames and J. D. Durtschi, *Clin. Chem.*, 2009, **55**, 641–658.

75    R. Stoltenburg and B. Strehlitz, *Int. J. Mol. Sci.*, 2018, **19**, 642.

76    Q. Hu, R. Wang, H. Wang, M. F. Slavik and Y. Li, *Sensors Actuators, B Chem.*, 2018, **273**, 220–227.

77    S. Kraemer, J. D. Vaught, C. Bock, L. Gold, E. Katilius, T. R. Keeney, N. Kim, N. A. Saccomano, S. K. Wilcox, D.

Zichi and G. M. Sanders, *PLoS One*, 2011, **6**, e26332.

78     J. C. Rohloff, A. D. Gelinas, T. C. Jarvis, U. A. Ochsner, D. J. Schneider, L. Gold and N. Janjic, *Mol. Ther. - Nucleic Acids*, 2014, **3**, e201.

79     Z. Zhuo, Y. Yu, M. Wang, J. Li, Z. Zhang, J. Liu, X. Wu, A. Lu, G. Zhang and B. Zhang, *Int. J. Mol. Sci.*, 2017, **18**, 1–19.

80     C. Tuerk and L. Gold, *Science (80-. ).*, 1990, **249**, 505–510.

81     J. Dagle, D. Weeks and  j Walder, 1991, **20**, 11–20.

82     G. M. Lamm, B. J. Blencowe, B. S. Sparoat, A. M. Iribarren, U. Ryder and A. I. Lamond, *Nucleic Acids Res.*, 1991, **19**, 3193–3198.

83     Z. Cao, R. Tong, A. Mishra, W. Xu, G. C. L. Wong, J. Cheng and Y. Lu, *Angew. Chemie - Int. Ed.*, 2009, **48**, 6494–6498.

84     C. S. M. Ferreira, M. C. Cheung, S. Missailidis, S. Bisland and J. Gariépy, *Nucleic Acids Res.*, 2009, **37**, 866–876.

85     Z. Bognár and R. E. Gyurcsányi, *Int. J. Mol. Sci.*, 2020, **21**, 1–40.

86     C. Chandola, S. Kalme, M. G. Casteleijn, A. Urtti and M. Neerathilingam, *J. Biosci.*, 2016, **41**, 535–561.

87     Aptamers vs antibodies, http://www.aptamergroup.co.uk/aptamers/aptamer-vs-antibodies/.

88     A. D. Keefe, S. Pai and A. Ellington, *Nat. Rev. Drug Discov.*, 2010.

89     P. Sundaram, H. Kurniawan, M. E. Byrne and J. Wower, *Eur. J. Pharm. Sci.*, 2013, **48**, 259–271.

90     D. M. Ecker, S. D. Jones and H. L. Levine, *MAbs*, 2014, **7**, 9–14.

91     Centauri Therapeutics, http://www.centauritherapeutics.com/.

92     A. Parashar, *J. Clin. Diagnostic Res.*, 2016, **10**, BE01–BE06.

93     S. A. Vinores, *Int. J. Nanomedicine*, 2006, **1**, 263–268.

94     D. J. Pieramici and M. D. Rabena, *Eye*, 2008, **22**, 1330–1336.

95     European Medicines Agency, 2012, **44**, 2010–2012.

96      A. H. C. Fong and T. Y. Y. Lai, *Clin. Interv. Aging*, 2013, **8**, 467–483.

97      A. Silman, *Br. J. Rheumatol.*, 1993, **32**, 89–91.

98      Y. Chernajovsky and A. Nissim, *Handb Exp Pharmacol.*, 2008, **181**, 3–18.

99      C. Chen, Z. Yang and X. Tang, *Med. Res. Rev.*, 2018, **38**, 829–869.

100     Alicaforsen, https://www.atlantichc.com/alicaforsen.

101     S. Kruspe, F. Mittelberger, K. Szameit and U. Hahn, *ChemMedChem*, 2014, **9**, 1998–2011.

102     W. C. Olson, W. D. W. Heston and A. K. Rajasekaran, *Rev. Recent Clin. Trials*, 2007, **2**, 182–190.

103     C. Liang, B. Guo, H. Wu, N. Shao, D. Li, J. Liu, L. Dang, C. Wang, H. Li, S. Li, W. K. Lau, Y. Cao, Z. Yang, C. Lu, X. He, D. W. T. Au, X. Pan, B. T. Zhang, C. Lu, H. Zhang, K. Yue, A. Qian, P. Shang, J. Xu, L. Xiao, Z. Bian, W. Tan, Z. Liang, F. He, L. Zhang, A. Lu and G. Zhang, *Nat. Med.*, 2015, **21**, 288–294.

104     D. C. Douek, M. Roederer and R. A. Koup, *Annu. Rev. Med.*, 2009, **60**, 471–484.

105     J. Zhou, P. Swiderski, H. Li, J. Zhang, C. P. Neff, R. Akkina and J. J. Rossi, *Nucleic Acids Res.*, 2009, **37**, 3094–3109.

106     J. Zhou, C. P. Neff, P. Swiderski, H. Li, D. D. Smith, T. Aboellail, L. Remling-Mulder, R. Akkina and J. J. Rossi, *Mol. Ther.*, 2013, **21**, 192–200.

107     P. Zhang, N. Zhao, Z. Zeng, Y. Feng, C. H. Tung, C. C. Chang and Y. Zu, *Lab. Investig.*, 2009, **89**, 1423–1432.

108     C. H. Weng, C. J. Huang and G. Bin Lee, *Sensors (Switzerland)*, 2012, **12**, 9514–9529.

109     B. P. Gray, M. D. Requena, M. D. Nichols and B. A. Sullenger, *Cell Chem. Biol.*, 2020, **27**, 232-244.e7.

110     J. H. Lee, M. V. Yigit, D. Mazumdar and Y. Lu, *Adv. Drug Deliv. Rev.*, 2010, **62**, 592–605.

111     D. Ferreira, J. Barbosa, D. A. Sousa, C. Silva, L. D. R. Melo, M. Avci-Adali, H. P. Wendel and L. R. Rodrigues, *Sci. Rep.*, 2021, **11**, 1–15.

112     Y. Wang, Y. Luo, T. Bing, Z. Chen, M. Lu, N. Zhang, D. Shangguan and X. Gao, *PLoS One*, 2014, **9**, 1–10.

113     Y. Song, Z. Zhu, Y. An, W. Zhang, H. Zhang, D. Liu, C. Yu, W. Duan and C. J. Yang, *Anal. Chem.*, 2013, **85**, 4141–4149.

114     P. Jolly, N. Formisano and P. Estrela, *Chem. Pap.*, 2015, **69**, 77–89.

115     Y. Tan, Y. Shi, X. Wu, H. Liang, Y. Gao, S. Li, X. Zhang, F. Wang and T. Gao, *Acta Pharmacol. Sin.*, 2013, **34**, 1491–1498.

116     P. Hong, W. Li and J. Li, *Sensors*, 2012, **12**, 1181–1193.

117     J. G. Bruno and J. L. Kiel, *Biosens. Bioelectron.*, 1999, **14**, 457–464.

118     X. Yan, X. Gao and Z. Zhang, *Genomics, proteomics Bioinforma. / Beijing Genomics Inst.*, 2004, **2**, 32–42.

119     M. Kuwahara and N. Sugimoto, *Molecules*, 2010, **15**, 5423–5444.

120     N. Derbyshire, S. J. White, D. H. J. Bunka, L. Song, S. Stead, J. Tarbin, M. Sharman, D. Zhou and P. G. Stockley, *Anal. Chem.*, 2012, **84**, 6595–6602.

121     M. Sioud, *RNA Interference*, Springer New York, New York, NY, 2015, vol. 1218.

122     G. Pasut and F. M. Veronese, *J. Control. Release*, 2012, **161**, 461–472.

123     P. Milla, F. Dosio and L. Cattel, *Curr. Drug Metab.*, 2011, **13**, 105–119.

124     L. Tan, K. G. Neoh, E. T. Kang, W. S. Choe and X. Su, *Macromol. Biosci.*, 2011, **11**, 1331–1335.

125     K. W. Thiel and P. H. Giangrande, *Oligonucleotides*, 2009, **19**, 209–222.

126     S. M. Nimjee, T. J. Povsic, B. A. Sullenger and R. C. Becker, *Nucleic Acid Ther.*, 2016, **26**, 147–155.

127     C. P. Rusconi, E. Scardino, J. Layzer, G. A. Pitoc, T. L. Ortel, D. Monroe and B. A. Sullenger, *Nature*, 2002, **419**, 90–94.

128     B. A. Sullenger, *J Thromb Haemost*, 2012, **78**, 711–716.

129     A. Michienzi, S. Li, J. A. Zaia and J. J. Rossi, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 14047–14052.

130     D. Ausländer, M. Wieland, S. Ausländer, M. Tigges and M. Fussenegger, *Nucleic Acids Res.*, 2011, **39**, 155.

131     L. Chaloin, M. J. Lehmann, G. Sczakiel and T. Restle, *Nucleic Acids Res.*, 2002, **30**, 4001–4008.

132     T. C. Chu, K. Y. Twu, A. D. Ellington and M. Levy, *Nucleic Acids Res.*, 2006, **34**, 1–6.

133     G. Mayer, M. Blind, W. Nagel, T. Böhm, T. Knorr, C. L. Jackson, W. Kolanus and M. Famulok, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 4961–4965.

134     E. Bertrand, L. Giver and A. Ellington, *Gene Ther.*, 1997, **4**, 45–54.

135 A. S. Davydova, M. A. Vorobjeva and A. G. Venyaminova, *Acta Naturae*, 2011, **3**, 12–29.

136 C. Meyer, K. Eydeler, E. Magbanua, T. Zivkovic, N. Piganeau, I. Lorenzen, J. Grötzinger, G. Mayer, S. Rose-John and U. Hahn, *RNA Biol.*, 2012, **9**, 67–80.

137 Y. Zhao, S. Duan, X. Zeng, C. Liu, N. M. Davies, B. Li and M. L. Forrest, *Mol. Pharm.*, 2012, **9**, 1705–1716.

138 K. M. Song, S. Lee and C. Ban, *Sensors*, 2012, **12**, 612–631.

139 L. V. Gening, S. A. Klincheva, A. Reshetnjak, A. P. Grollman and H. Miller, *Nucleic Acids Res.*, 2006, **34**, 2579–2586.

140 A. V. Lakhin, A. A. Kazakov, A. V. Makarova, Y. I. Pavlov, A. S. Efremova, S. I. Shram, V. Z. Tarantul and L. V. Gening, *Nucleic Acid Ther.*, 2012, **22**, 49–57.

141 J. C. Cox and A. D. Ellington, *Bioorganic Med. Chem.*, 2001, **9**, 2525–2531.

142 D. Eulberg, K. Buchner, C. Maasch and S. Klussmann, *Nucleic Acids Res.*, 2005, **33**, 1–10.

143 M. Berezovski, A. Drabovich, S. M. Krylova, M. Musheev, V. Okhonin, A. Petrov and S. N. Krylov, *J. Am. Chem. Soc.*, 2005, **127**, 3165–3171.

144 A. P. Drabovich, M. Berezovski, V. Okhonin and S. N. Krylov, *Anal. Chem.*, 2006, **78**, 3171–3178.

145 S. Gupta, D. W. Drolet, S. K. Wolk, S. M. Waugh, J. C. Rohloff, J. D. Carter, W. S. Mayfield, M. R. Otis, C. R. Fowler, T. Suzuki, M. Hirota, Y. Ishikawa, D. J. Schneider and N. Janjic, *Nucleic Acid Ther.*, 2017, **27**, 683.

146 D. H. J. Bunka, O. Platonova and P. G. Stockley, *Curr. Opin. Pharmacol.*, 2010, **10**, 557–562.

147 J. Ouellet, *Front. Chem.*, 2016, **4**, 1–12.

148 A. Kertsburg and G. A. Soukup, *Nucleic Acids Res.*, 2002, **30**, 4599–4606.

149 A. D. Ellington and J. W. Szostak, *Nature*, 1990, **346**, 818–822.

150 D. L. Robertson and G. F. Joyce, *Lett. to Nat.*, 1990, **344**, 467–468.

151 S. L. Paige, S. Thomas, C. L. Stoick-cooper, H. Wang, R. Sandstrom, L. Pabon, H. Reinecke, G. Pratt, R. T. Moon, J. Stamatoyannopoulos and C. E. Murry, 2013, **151**, 221–232.

152 N. J. N. N. P. S. Daniel E Shumer, *Physiol. Behav.*, 2017, **176**, 139–148.

153 Avula, *Curr Opin Chem Biol.*, 2014, **23**, 1–7.

154     D. P. I. Gil and C. Clayton, *PLoS One*, 2016, **11**, 1–13.

155     D. Vitiello, D. B. Pecchia and J. M. Burke, *Rna*, 2000, **6**, 628–637.

156     D. Musumeci, C. Platella, C. Riccardi, F. Moccia and D. Montesarchio, *Cancers (Basel).*, 2017, **9**, 1–43.

157     Z. Chen, P. A. Lichtor, A. P. Berliner, J. C. Chen and D. R. Liu, *Nat. Chem.*, 2018, **10**, 420–427.

158     G. E. Lee, J. Kim, J. S. Lee, J. S. Ko, E. J. Lee and J. S. Yoon, *Front. Endocrinol. (Lausanne).*, 2021, **11**, 1–11.

159     B. N. Gawande, J. C. Rohloff, J. D. Carter, I. von Carlowitz, C. Zhang, D. J. Schneider and N. Janjic, *Proc. Natl. Acad. Sci.*, 2017, **114**, 2898–2903.

160     E. Cereda, A. Brambilla, A. Ezhaya, W. Gaida, K. Mendla, E. Monferini, N. Pitsikas, G. B. Schiavi and A. Sagrada, *Eur. J. Drug Metab. Pharmacokinet.*, 1994, **19**, 179–183.

161     K. S. Lam, M. Lebl and V. Krchňák, *Chem. Rev.*, 1997, **97**, 411–448.

162     R. A. Houghten, C. Pinilla, S. E. Blondelle, J. R. Appel, C. T. Dooley and J. H. Cuervo, 1991, **354**, 19–21.

163     H. M. Geysen, R. H. Meloen and S. J. Barteling, *Proc. Natl. Acad. Sci. U. S. A.*, 1984, **81**, 3998–4002.

164     K. S. Lam, S. E. Salmon, E. M. Hersh, V. J. Hruby, W. M. Kazmierski and R. J. Knapp, *Nature*, 1991, **354**, 82–84.

165     Barry A. Bunin and Jonathan A. Ellman, *J. Am. Chem. Soc.*, 1992, **775**, 10997–10998.

166     G. P. Smith, *Science*, 1985, **228**, 1315–1317.

167     1 Adam Frankel, 1 Steven W. Millward and  and R. W. Roberts, *Chem. Biol.*, 2003, **10**, 1043–1050.

168     K. Josephson, M. C. T. Hartman and J. W. Szostak, *J. Am. Chem. Soc.*, 2005, **127**, 11727–11735.

169     H. Murakami, A. Ohta, H. Ashigai and H. Suga, *Nat. Methods*, 2006, **3**, 357–359.

170     C. Heinis, T. Rutherford, S. Freund and G. Winter, *Nat. Chem. Biol.*, 2009, **5**, 502–507.

171     R. Liu, X. Li and K. S. Lam, *Curr. Opin. Chem. Biol.*, 2017, **38**, 117–126.

172     R. A. Goodnow, C. E. Dumelin and A. D. Keefe, *Nat. Rev. Drug Discov.*, 2017, **16**, 131–147.

173     S. Brenner and R. A. Lerner, *Proc. Natl. Acad. Sci. U. S. A.*, 1992, **89**, 5381–5383.

174     A. Martín, C. A. Nicolaou and M. A. Toledo, *Commun. Chem.*, 2020, **3**, 1–9.

175     D. Neri and R. A. Lerner, *Annu. Rev. Biochem.*, 2018, **87**, 479–502.

176     V. Kunig, M. Potowski, A. Gohla and A. Brunschweiger, *Biol. Chem.*, 2018, **399**, 691–710.

177     M. Song and G. T. Hwang, *J. Med. Chem.*, 2020, **63**, 6578–6599.

178     G. Zhao, Y. Huang, Y. Zhou, Y. Li and X. Li, *Expert Opin. Drug Discov.*, 2019, **14**, 735–753.

179     R. Liu, J. Marik and K. S. Lam, *J Am Chem Soc*, 2002, **124**, 7678–80.

180     M. Meldal, *Curr. Opin. Chem. Biol.*, 2004, **8**, 238–244.

181     K. D. Foust and B. K. Kaspar, 2010, **8**, 4017–4018.

182     W. Xiao, Y. Wang, E. Y. Lau, J. Luo, N. Yao, C. Shi, L. Meza, H. Tseng, Y. Maeda, P. Kumaresan, R. Liu, F. C. Lightstone, Y. Takada and K. S. Lam, *Mol. Cancer Ther.*, 2010, **9**, 2714–2723.

183     A. J. Quartararo, Z. P. Gates, B. A. Somsen, N. Hartrampf, X. Ye, A. Shimada, Y. Kajihara, C. Ottmann and B. L. Pentelute, *Nat. Commun.*, 2020, **11**, 1–11.

184     T. B. Trinh, Q. Xiao and D. Pei, *Biochemistry*, 2013, **52**, 5645–5655.

185     M. N. Idso, A. S. Akhade, M. L. Arrieta-Ortiz, B. T. Lai, V. Srinivas, J. P. Hopkins, A. O. Gomes, N. Subramanian, N. Baliga and J. R. Heath, *Chem. Sci.*, 2020, **11**, 3054–3067.

186     K. Sriram and P. A. Insel, *Mol. Pharmacol.*, 2018, **93**, 251–258.

187     B. P. Gray and K. C. Brown, *Chem. Rev.*, 2014, **114**, 1020–1081.

188     Goate and APP, *Lett. To Nat.*, 1991, **353**, 737–740.

189     X. Chen and S. S. Gambhir, *Nat. Chem. Biol.*, 2006, **2**, 351–352.

190     H. Jin and J. Varner, *Br. J. Cancer*, 2004, **90**, 561–565.

191     W. Xiao, Y. Wang, E. Y. Lau, J. Luo, N. Yao, C. Shi, L. Meza, H. Tseng, Y. Maeda, P. Kumaresan, R. Liu, F. C. Lightstone, Y. Takada and K. S. Lam, *Mol. Cancer Ther.*, 2010, **9**, 2714–2723.

192     O. Erharuyi, S. Simanski, P. J. McEnaney and T. Kodadek, *Bioorganic Med. Chem. Lett.*, 2018, **28**, 2773–2778.

193     K. S. Anderson and J. LaBaer, *J. Proteome Res.*, 2005, **4**, 1123–1133.

194    T. M. Doran, M. Sarkar and T. Kodadek, *J. Am. Chem. Soc.*, 2016, **138**, 6076–6094.

195    T. M. Doran, S. Simanski and T. Kodadek, *ACS Chem. Biol.*, 2015, **10**, 401–412.

196    P. Blakskjaer, T. Heitner and N. J. V. Hansen, *Curr. Opin. Chem. Biol.*, 2015, **26**, 62–71.

197    F. Samain, T. Ekblad, G. Mikutis, N. Zhong, M. Zimmermann, A. Nauer, D. Bajic, W. Decurtins, J. Scheuermann, P. J. Brown, J. Hall, S. Gräslund, H. Schüler, D. Neri and R. M. Franzini, *J. Med. Chem.*, 2015, **58**, 5143–5149.

198    J. P. Daguer, C. Zambaldo, M. Ciobanu, P. Morieux, S. Barluenga and N. Winssinger, *Chem. Sci.*, 2015, **6**, 739–744.

199    R. Mamidala, P. Majumdar, K. K. Jha, C. Bathula, R. Agarwal, M. T. Chary, H. K. Mazumdar, P. Munshi and S. Sen, *Sci. Rep.*, 2016, **6**, 1–13.

200    C. Aquino, M. Sarkar, M. J. Chalmers, K. Mendes, T. Kodadek and G. C. Micalizio, *Nat. Chem.*, 2012, **4**, 99–104.

201    M. Guan, W. Yao, R. Liu, K. S. Lam, J. Nolta, J. Jia, B. Panganiban, L. Meng, P. Zhou, M. Shahnazari, R. O. Ritchie and N. E. Lane, *Nat. Med.*, 2012, **18**, 456–462.

202    A. E. Corson, S. A. Armstrong, M. E. Wright, E. E. McClelland and K. L. Bicker, *ACS Med. Chem. Lett.*, 2016, **7**, 1139–1144.

203    T. B. Trinh, P. Upadhyaya, Z. Qian and D. Pei, *ACS Comb. Sci.*, 2016, **18**, 75–85.

204    R. Fleeman, T. M. Lavoi, R. G. Santos, A. Morales, A. Nefzi, G. S. Welmaker, J. L. Medina-Franco, M. A. Giulianotti, R. A. Houghten and L. N. Shaw, *J. Med. Chem.*, 2015, **58**, 3340–3355.

205    J. Baran, A. Pituch-Noworolska, A. Krzeszowiak, J. Wieckiewicz, J. Stachura, J. Pryjma, T. Popiela, A. Szczepanik and M. Zembala, *Int. J. Mol. Med.*, 1998, **1**, 573–578.

206    J. Zimmermann, *Eur. J. Immunol.*, 2017, **47**, 1584–1797.

207    J. J. Mcsharry, *Clin. Microbiol. Rev.*, 1994, **7**, 576–604.

208    G. Janossy and H. Shapiro, *Cytom. Part B - Clin. Cytom.*, 2008, **74**, 6–10.

209    B. H. Villas, *Cell Vis.*, 1998, **5**, 56–61.

210    Flourescent Activiated Cell Sorter, http://www.sinobiological.com/flow-cytometry-fcm-facs-fluorescence-activated-cell-sorting-facs.html (Accessed 26/09/2018.

211    A. Manuscript, 2013, **402**, 325–335.

212    D. I. Strokotov, M. A. Yurkin, K. V. Gilev, D. R. van Bockstaele, A. G. Hoekstra, N. B. Rubtsov and V. P. Maltsev, *J. Biomed. Opt.*, 2009, **14**, 064036.

213    C. E. Stanciu, Y. J. Kwon and C. J. Ehrhardt, *Data Br.*, 2016, **6**, 416–418.

214    S. C. Pruitt, L. M. Mielnicki and C. C. Stewart, *Methods Mol. Biol.*, 2004, **263**, 239–258.

215    B. Dickinson, *BD Influx ™ Cell Sorter User ' s Guid.*, 2011, 1–378.

216    L. A. Herzenberg, J. Tung, W. A. Moore, L. A. Herzenberg and D. R. Parks, *Nat. Immunol.*, 2006, **7**, 681–685.

217    J. A. Lee, J. Spidlen, K. Boyce, J. Cai, N. Crosbie, M. Dalphin, J. Furlong, M. Gasparetto, M. Goldberg, E. M. Goralczyk, B. Hyun, K. Jansen, T. Kollmann, M. Kong, R. Leif, S. McWeeney, T. D. Moloshok, W. Moore, G. Nolan, J. Nolan, J. Nikolich-Zugich, D. Parrish, B. Purcell, Y. Qian, B. Selvaraj, C. Smith, O. Tchuvatkina, A. Wertheimer, P. Wilkinson, C. Wilson, J. Wood, R. Zigon, R. H. Scheuermann and R. R. Brinkman, *Cytom. Part A*, 2008, **73**, 926–930.

218    K. Kang, S. B. Lee, J. H. Yoo and C. W. Nho, *Biotechnol. Lett.*, 2010, **32**, 1045–1052.

219    Bruce S. Edwards and Larry A. Sklar, *J. Biomol. Screen.*, 2015, **1**, 1–19.

220    M. Ding, K. Kaspersson, D. Murray and C. Bardelle, *Drug Discov. Today*, 2017, **22**, 1844–1850.

221    J. P. Robinson, V. Patsekin, B. Rajwa, V. Jo Davisson, N. Li and P. Narayanan, *Genet. Eng. Biotechnol. News*, 2011, **31**, 20–22.

222    J. Chen, S. M. Young, C. Allen, A. Seeber, M. P. Péli-Gulli, N. Panchaud, A. Waller, O. Ursu, T. Yao, J. E. Golden, J. J. Strouse, M. B. Carter, H. Kang, C. G. Bologa, T. D. Foutz, B. S. Edwards, B. R. Peterson, J. Aubé, M. Werner-Washburne, R. J. Loewith, C. De Virgilio and L. A. Sklar, *ACS Chem. Biol.*, 2012, **7**, 715–722.

223    S. M. Young, C. Bologa, E. R. Prossnitz, T. I. Oprea, L. A. Sklar and B. S. Edwards, *J. Biomol. Screen.*, 2005, **10**, 374–382.

224    Z. Zhao, M. K. Haynes, O. Ursu, B. S. Edwards, L. A. Sklar and A. Zweifach, *J. Biomol. Screen.*, 2015, **20**, 359–371.

225 R. F. Curpan, P. C. Simons, D. Zhai, S. M. Young, M. B. Carter, C. G. Bologa, T. I. Oprea, A. C. Satterthwait, J. C. Reed, B. S. Edwards and L. A. Sklar, *Assay Drug Dev. Technol.*, 2011, **9**, 465–474.

226 D. L. Roman, J. N. Talbot, R. A. Roof, R. K. Sunahara, J. R. Traynor and R. R. Neubig, *Mol. Pharmacol.*, 2007, **71**, 169–175.

227 Z. Surviladze, S. M. Young and L. A. Sklar, *Methods Mol. Biol.*, 2012, **827**, 253–270.

228 S. M. Young, C. M. Bologa, D. Fara, B. K. Bryant, J. J. Strouse, J. B. Arterburn, R. D. Ye, T. I. Oprea, E. R. Prossnitz, L. A. Sklar and B. S. Edwards, *Cytom. Part A*, 2009, **75**, 253–263.

229 D. Cook, D. Brown, R. Alexander, R. March, P. Morgan, G. Satterthwaite and M. N. Pangalos, *Nat. Rev. Drug Discov.*, 2014, **13**, 419–431.

230 J. Neubauer, *J. Appl. Physiol.*, 2001, **90**, 1593–1599.

231 Asthma Uk, https://www.asthma.org.uk/about/media/facts-and-statistics/.

232 World Health Organisation, https://www.who.int/news-room/fact-sheets/detail/asthma.

233 V. Sauzeau, M. A. Sevilla, J. V. Rivas-Elena, E. De Álava, M. J. Montero, J. M. López-Novoa and X. R. Bustelo, *Nat. Med.*, 2006, **12**, 841–845.

234 Z. Surviladze, A. Waller, Y. Wu, E. Romero, B. S. Edwards, A. Wandinger-Ness and L. A. Sklar, *J. Biomol. Screen.*, 2010, **15**, 10–20.

235 D. Wu, C. K. L. Gordon, J. H. Shin, M. Eisenstein and H. T. Soh, *Acc. Chem. Res.*, 2022, **55**, 685–695.

236 G. van den E. Sherrif F. Ibrahim, *Biotechniques*, 2007, **56**, 19–39.

237 F. E. Craig and K. A. Foon, *Blood*, 2008, **111**, 3941–3967.

238 A. Wood, A. Shukla, J. Schneider, J. S. Lee, J. D. Stanton, T. Dzuiba, S. K. Swanson, L. Florens, M. P. Washburn, J. Wyrick, S. R. Bhaumik and A. Shilatifard, *Mol. Cell. Biol.*, 2007, **27**, 709–720.

239 B. Barlogie, M. N. Raber, J. Schumann, T. S. Johnson, B. Drewinko, D. E. Swartzendruber, W. Göhde, M. Andreeff, E. J. Freireich, B. Barlogie, M. N. Raber, J. Schumann, T. S. Johnson, B. Drewinko, D. E. Swartzendruber, W. Göhde, M. Andreeff, E. J. Freireich, B. Barlogie, M. N. Raber, J. Schumann, T. S. Johnson, B. Drewinko, D. E. Swartzendruber, W. Göhde, M. Andreeff, E. J. Freireich, B. Barlogie, M. N. Raber, J. Schumann, T. S. Johnson, B. Drewinko, D. E. Swartzendruber, W. Göhde, M. Andreeff and E. J.

Freireich, *Cancer Res.*, 1983, **43**, 3982–3997.

240     M. L. Friedlander, D. W. Hedley and I. W. Taylor, *J. Clin. Pathol.*, 1984, **37**, 961–974.

241     D. E. Merkel, L. G. Dressler and W. L. McGuire, *J. Clin. Oncol.*, 1987, **5**, 1690–1703.

242     P. M. Thompson, K. M. Hayashi, G. I. De Zubicaray, A. L. Janke, S. E. Rose, J. Semple, M. S. Hong, D. H. Herman, D. Gravano, D. M. Doddrell and A. W. Toga, *Neuroimage*, 2004, **22**, 1754–1766.

243     S. A. Motzer, M. Jarrett, M. M. Heitkemper and J. Tsuji, *Biol. Res. Nurs.*, 2002, **4**, 31–42.

244     S. Gennaro, W. Fehder, P. Gallagher, S. Miller, S. D. Douglas and D. E. Campbell, *Clin. Diagn. Lab. Immunol.*, 1997, **4**, 195–201.

245     G. Carioli, M. Malvezzi, P. Bertuccio, P. Boffetta, F. Levi, C. La Vecchia and E. Negri, *Ann. Oncol.*, 2021, **32**, 478–487.

246     R. L. Siegel, K. D. Miller, H. E. Fuchs and A. Jemal, *CA. Cancer J. Clin.*, 2021, **71**, 7–33.

247     R. Gosain, Y. Abdou, A. Singh, N. Rana, I. Puzanov and M. S. Ernstoff, *Curr. Oncol. Rep.*, 2020, **22**, 53.

248     A. G. Lai, L. Pasea, A. Banerjee, G. Hall, S. Denaxas, W. H. Chang, M. Katsoulis, B. Williams, D. Pillay, M. Noursadeghi, D. Linch, D. Hughes, M. D. Forster, C. Turnbull, N. K. Fitzpatrick, K. Boyd, G. R. Foster, T. Enver, V. Nafilyan, B. Humberstone, R. D. Neal, M. Cooper, M. Jones, K. Pritchard-Jones, R. Sullivan, C. Davie, M. Lawler and H. Hemingway, *BMJ Open*, 2020, **10**, 1–9.

249     P. A. Wingo, C. J. Cardinez, S. H. Landis, R. T. Greenlee, L. A. G. Ries, R. N. Anderson and M. J. Thun, *Cancer*, 2003, **97**, 3133–3275.

250     National Cancer Institute, https://www.cancer.gov/about-cancer/understanding/.

251     R. S. Herbst, *Int. J. Radiat. Oncol. Biol. Phys.*, 2004, **59**, 21–26.

252     B. Alberts, D. Bray, K. Hopkin, A. Johnson, J. Lewis, M. Raff, K. Roberts and P. Walter, *Essential Cell Biology*, 2014.

253     D. X. Nguyen, P. D. Bos and J. Massagué, *Nat. Rev. Cancer*, 2009, **9**, 274–284.

254     D. Hanahan and R. A. Weinberg, *Cell*, 2000, **100**, 57–70.

255     D. Hanahan and R. A. Weinberg, *Cell*, 2011, **144**, 646–674.

256     S. C. Sfdfqupst, G. Podphfoft, H. Zhang, A. Berezov, Q. Wang, G. Zhang, J. Drebin, R. Murali, M. I. Greene, A. Gschwind, O. M. Fischer, A. Ullrich, S. Yuzawa, Y. Opatowsky, Z. Zhang, V. Mandiyan, I. Lax, J. Schlessinger and R. Roskoski, *Cell*, 2007, **79**, 323–334.

257     J. Mendelsohn and J. Baselga, *Semin. Oncol.*, 2006, **33**, 369–385.

258     M. A. Lemmon and J. Schlessinger, *Cell*, 2011, **141**, 1117–1134.

259     H. W. Huang, S. K. Mohan and C. Yu, *Biochem. Biophys. Res. Commun.*, 2010, **402**, 705–710.

260     R. J. Bodnar, *Adv. Wound Care*, 2013, **2**, 24–29.

261     J. Chen, F. Zeng, S. J. Forrester, S. Eguchi, M. Z. Zhang and R. C. Harris, *Physiol. Rev.*, 2016, **96**, 1025–1069.

262     J. R. Grandis and J. C. Sok, *Pharmacol. Ther.*, 2004.

263     A. Wells, *Int. J. Biochem. Cell Biol.*, 1999, **31**, 637–643.

264     M. A. Olayioye, *EMBO J.*, 2000, **19**, 3159–3167.

265     K. M. Ferguson, M. B. Berger, J. M. Mendrola, H.-S. Cho, D. J. Leahy and M. A. Lemmon, *Mol. Cell*, 2003, **11**, 507–517.

266     H. Ogiso, R. Ishitani, O. Nureki, S. Fukai, M. Yamanaka, J. H. Kim, K. Saito, A. Sakamoto, M. Inoue, M. Shirouzu and S. Yokoyama, *Cell*, 2002, **110**, 775–787.

267     H. S. Cho and D. J. Leahy, *Science (80-. ).*, 2002, **297**, 1330–1333.

268     T. P. J. Garrett, N. M. McKern, M. Lou, T. C. Elleman, T. E. Adams, G. O. Lovrecz, H. J. Zhu, F. Walker, M. J. Frenkel, P. A. Hoyne, R. N. Jorissen, E. C. Nice, A. W. Burgess and C. W. Ward, *Cell*, 2002, **110**, 763–773.

269     S. Bouyain, P. A. Longo, S. Li, K. M. Ferguson and D. J. Leahy, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 15024–15029.

270     Y. Yarden and M. X. Sliwkowski, *Nat. Rev. Mol. Cell Biol.*, 2001, **2**, 127–137.

271     A. Citri and Y. Yarden, *Nat. Rev. Mol. Cell Biol.*, 2006, **7**, 505–516.

272     T. Hunter, *Cold Spring Harb. Perspect. Biol.*, 2014, **6**, a020644–a020644.

273     D. R. Siwak, M. Carey, B. T. Hennessy, C. T. Nguyen, M. J. McGahren Murray, L. Nolden and G. B. Mills, *J. Oncol.*, 2010, **2010**, 1–20.

389

274    E. Tzahar, H. Waterman, X. Chen, G. Levkowitz, D. Karunagaran, S. Lavi, B. J. Ratzkin and Y. Yarden, *Mol. Cell. Biol.*, 1996, **16**, 5276–5287.

275    J. Baselga and J. Albanell, *Hematol. Oncol. Clin. North Am.*, 2002, **16**, 1041–1063.

276    C. Sweeney and K. L. Carraway, *Oncogene*, 2000, **19**, 5568–5573.

277    N. Prenzel, O. M. Fischer, S. Streit, S. Hart and A. Ullrich, *Endocr. Relat. Cancer*, 2001, **8**, 11–31.

278    L. Poitras, S. Jean, N. Islam and T. Moss, *FEBS Lett.*, 2003, **543**, 129–135.

279    D. D. Hirsch and P. J. S. Stork, *J. Biol. Chem.*, 1997, **272**, 4568–4575.

280    T. Fukazawa, S. Miyake, V. Band and H. Band, *J. Biol. Chem.*, 1996, **271**, 14554–14559.

281    R. Katso, K. Okkenhaug, K. Ahmadi, J. Timms and M. D. Waterfield, *Annu. Rev. Cell Dev. Biol.*, 2001, **17**, 615–75.

282    S. Brader and S. A. Eccles, *Tumori*, 2004, **90**, 2–8.

283    C. Tkaczyk, M. A. Beaven, S. M. Brachman, D. D. Metcalfe and A. M. Gilfillan, *J. Biol. Chem.*, 2003, **278**, 48474–48484.

284    M. Hamilton, J. Liao, M. K. Cathcart and A. Wolfman, *J. Biol. Chem.*, 2001, **276**, 29079–29090.

285    K. C. Corbit, N. Trakul, E. M. Eves, B. Diaz, M. Marshall and M. R. Rosner, *J. Biol. Chem.*, 2003, **278**, 13061–13068.

286    A. Altman and M. Villalba, *Immunol. Rev.*, 2003, **192**, 53–63.

287    C. M. Silva, *Oncogene*, 2004, **23**, 8017–8023.

288    H. W. Lo, S. C. Hsu, W. Xia, X. Cao, J. Y. Shih, Y. Wei, J. L. Abbruzzese, G. N. Hortobagyi and M. C. Hung, *Cancer Res.*, 2007, **67**, 9066–9076.

289    S. Barr, S. Thomson, E. Buck, S. Russo, F. Petti, I. Sujka-Kwok, A. Eyzaguirre, M. Rosenfeld-Franklin, N. W. Gibson, M. Miglarese, D. Epstein, K. K. Iwata and J. D. Haley, *Clin. Exp. Metastasis*, 2008, **25**, 685–693.

290    M. Grände, Å. Franzen, J. O. Karlsson, L. E. Ericson, N. E. Heldin and M. Nilsson, *J. Cell Sci.*, 2002, **115**, 4227–4236.

291    N. G. Docherty, O. E. O'Sullivan, D. A. Healy, M. Murphy, A. J. O'Neill, J. M. Fitzpatrick and R. W. G. Watson,

*Am. J. Physiol. - Ren. Physiol.*, 2006, **290**, 1202–1213.

292    Y. C. Tian, Y. C. Chen, C. T. Chang, C. C. Hung, M. S. Wu, A. Phillips and C. W. Yang, *Exp. Cell Res.*, 2007, **313**, 2367–2377.

293    N. Ahmed, S. Maines-Bandiera, M. A. Quinn, W. G. Unger, S. Dedhar and N. Auersperg, *Am. J. Physiol. - Cell Physiol.*, 2006, **290**, 1532–1543.

294    F. Ciardiello, *Drugs*, 2000, **60**, 25–32.

295    F. Ciardiello and G. Tortora, *Eur. J. Cancer*, 2003, **39**, 1348–1354.

296    D. S. Salomon, R. Brandt, F. Ciardiello and N. Normanno, *Crit. Rev. Oncol. Hematol.*, 1995, **19**, 183–232.

297    J. G. M. Klijn, P. M. J. J. Berns, P. I. M. Schmitz and J. A. Foekens, *Endocr.Rev.*, 11AD, **Endocrine**-, 1–17.

298    S. Mabuchi, M. Ohmichi, A. Kimura, K. Hisamoto, J. Hayakawa, Y. Nishio, K. Adachi, K. Takahashi, E. Arimoto-Ishida, Y. Nakatsuji, K. Tasaka and Y. Murata, *J. Biol. Chem.*, 2002, **277**, 33490–33500.

299    M. Hatakeyama, S. Kimura, T. Naka, T. Kawasaki, N. Yumoto, M. Ichikawa, J. H. Kim, K. Saito, M. Saeki, M. Shirouzu, S. Yokoyama and A. Konagaya, *Biochem. J.*, 2003, **373**, 451–463.

300    M. A. Bareschino, C. Schettino, T. Troiani, E. Martinelli, F. Morgillo and F. Ciardiello, *Ann. Oncol.*, 2007, **18**, vi35–vi41.

301    A. W. Burgess, H. S. Cho, C. Eigenbrot, K. M. Ferguson, T. P. J. Garrett, D. J. Leahy, M. A. Lemmon, M. X. Sliwkowski, C. W. Ward and S. Yokoyama, *Mol. Cell*, 2003, **12**, 541–552.

302    H. Sunada, B. E. Magun, J. Mendelsohn and C. L. MacLeod, *Proc. Natl. Acad. Sci. U. S. A.*, 1986, **83**, 3825–3829.

303    H. Kimura, K. Sakai, T. Arao, T. Shimoyama, T. Tamura and K. Nishio, *Cancer Sci.*, 2007, **98**, 1275–1280.

304    B. W. Ennis, M. E. Lippman and R. B. Dickson, *Cancer Invest.*, 1991, **9**, 553–562.

305    J. H. J. M. Van Krieken, A. Jung, T. Kirchner, F. Carneiro, R. Seruca, F. T. Bosman, P. Quirke, J. F. Fléjou, T. Plato Hansen, G. De Hertogh, P. Jares, C. Langner, G. Hoefler, M. Ligtenberg, D. Tiniakos, S. Tejpar, G. Bevilacqua and A. Ensari, *Virchows Arch.*, 2008, **453**, 417–431.

306    A. K. Coutinho and C. M. S. Rocha Lima, *Cancer Control*, 2003, **10**, 224–238.

307     R. B. Cohen, *Clin. Colorectal Cancer*, 2003, **2**, 246–251.

308     D. Cunningham, Y. Humblet, S. Siena, D. Khayat, H. Bleiberg, A. Santoro, D. Bets, M. Mueser, A. Harstrick, C. Verslype, I. Chau and E. Van Cutsem, *N. Engl. J. Med.*, 2004, **351**, 337–345.

309     J. T. Thigpen, *Yearb. Oncol.*, 2011, **2011**, 199–201.

310     M. H. Cohen, H. Chen, S. Shord, C. Fuchs, K. He, H. Zhao, S. Sickafuse, P. Keegan and R. Pazdur, *Oncologist*, 2013, **18**, 460–466.

311     Medical Advisory Secretariat, *Ont. Health Technol. Assess. Ser.*, 2010, **10**, 1–49.

312     FDA approves new dosing regimen for cetuximab, https://www.fda.gov/drugs/resources-information-approved-drugs/fda-approves-new-dosing-regimen-cetuximab.

313     S. Li, K. R. Schmitz, P. D. Jeffrey, J. J. W. Wiltzius, P. Kussie and K. M. Ferguson, *Cancer Cell*, 2005, **7**, 301–311.

314     M. Hojjat-Farsangi, *Int. J. Mol. Sci.*, 2014, **15**, 13768–13801.

315     R. S. Herbst, M. Fukuoka and J. Baselga, *Nat. Rev. Cancer*, 2004, **4**, 956–965.

316     J. Dowell, J. D. Minna and P. Kirkpatrick, *Nat. Rev. Drug Discov.*, 2005, **4**, 13–14.

317     K. W. Robinson and A. B. Sandler, *Curr. Oncol. Rep.*, 2013, **15**, 396–404.

318     P. O'Donnell, J. Ferguson, J. Shyu, R. Current, T. Rehage, J. Tsai, M. Christensen, H. B. Tran, S. S. C. Chien, F. Shieh, W. Wei, H. J. Lawrence, L. Wu, R. Schilling, K. Bloom, W. Maltzman, S. Anderson and S. Soviero, *BMC Cancer*, 2013, **13**, 1–10.

319     M. J. Moore, D. Goldstein, J. Hamm, A. Figer, J. R. Hecht, S. Gallinger, H. J. Au, P. Murawa, D. Walde, R. A. Wolff, D. Campos, R. Lim, K. Ding, G. Clark, T. Voskoglou-Nomikos, M. Ptasynski and W. Parulekar, *J. Clin. Oncol.*, 2007, **25**, 1960–1966.

320     R. J. Schilder, M. W. Sill, Y. C. Lee and R. Mannel, *Int. J. Gynecol. Cancer*, 2009, **19**, 929–933.

321     Z. A. Wainberg, L. S. Lin, B. Dicarlo, K. M. Dao, R. Patel, D. J. Park, H. J. Wang, R. Elashoff, N. Ryba and J. R. Hecht, *Br. J. Cancer*, 2011, **105**, 760–765.

322     T. Dragovich, S. McCoy, C. M. Fenoglio-Preiser, J. Wang, J. K. Benedetti, A. F. Baker, C. B. Hackett, S. G. Urba, K. S. Zaner, C. D. Blanke and J. L. Abbruzzese, *J. Clin. Oncol.*, 2006, **24**, 4922–4927.

323    J. M. Llovet and V. Hernandez-Gea, *Clin. Cancer Res.*, 2014, **20**, 2072–2079.

324    M. S. Gordon, M. Hussey, R. B. Nagle, P. N. Lara, P. C. Mack, J. Dutcher, W. Samlowski, J. I. Clark, D. I. Quinn, C. X. Pan and D. Crawford, *J. Clin. Oncol.*, 2009, **27**, 5788–5793.

325    I. Sullivan and D. Planchard, *Front. Med.*, 2016, **3**, 1–13.

326    T. J. Lynch, D. W. Bell, R. Sordella, S. Gurubhagavatula, R. A. Okimoto, B. W. Brannigan, P. L. Harris, S. M. Haserlat, J. G. Supko, F. G. Haluska, D. N. Louis, D. C. Christiani, J. Settleman and D. A. Haber, *N. Engl. J. Med.*, 2004, **350**, 2129–2139.

327    W. Pao, V. Miller, M. Zakowski, J. Doherty, K. Politi, I. Sarkaria, B. Singh, R. Heelan, V. Rusch, L. Fulton, E. Mardis, D. Kupfer, R. Wilson, M. Kris and H. Varmus, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 13306–13311.

328    G. R. Oxnard, K. S. Thress, R. S. Alden, R. Lawrance, C. P. Paweletz, M. Cantarini, J. C. H. Yang, J. C. Barrett and P. A. Jänne, *J. Clin. Oncol.*, 2016, **34**, 3375–3382.

329    L. V. Sequist, B. Besse, T. J. Lynch, V. A. Miller, K. K. Wong, B. Gitlitz, K. Eaton, C. Zacharchuk, A. Freyman, C. Powell, R. Ananthakrishnan, S. Quinn and J. C. Soria, *J. Clin. Oncol.*, 2010, **28**, 3076–3083.

330    K. L. Reckamp, G. Giaccone, D. R. Camidge, S. M. Gadgeel, F. R. Khuri, J. A. Engelman, M. Koczywas, A. Rajan, A. K. Campbell, D. Gernhardt, A. Ruiz-Garcia, S. Letrent, J. Liang, I. Taylor, J. P. O'Connell and P. A. Jänne, *Cancer*, 2014, **120**, 1145–1154.

331    D. A. E. Cross, S. E. Ashton, S. Ghiorghiu, C. Eberlein, C. A. Nebhan, P. J. Spitzler, J. P. Orme, M. R. V. Finlay, R. A. Ward, M. J. Mellor, G. Hughes, A. Rahi, V. N. Jacobs, M. R. Brewer, E. Ichihara, J. Sun, H. Jin, P. Ballard, K. Al-Kadhimi, R. Rowlinson, T. Klinowska, G. H. P. Richmond, M. Cantarini, D. W. Kim, M. R. Ranson and W. Pao, *Cancer Discov.*, 2014, **4**, 1046–1061.

332    N. Karachaliou, M. Fernandez-Bruno, J. W. P. Bracht and R. Rosell, *Transl. Cancer Res.*, 2019, **8**, S23–S47.

333    J. Remon, C. E. Steuer, S. S. Ramalingam and E. Felip, *Ann. Oncol.*, 2018, **29**, i20–i27.

334    N. Li, J. N. Ebright, G. M. Stovall, X. Chen, H. H. Nguyen, A. Singh, A. Syrett and A. D. Ellington, *J. Proteome Res.*, 2009, **8**, 2438–2448.

335    J. M. Palomo, *RSC Adv.*, 2014, **4**, 32658–32672.

336     A. L. Marzinzik and E. R. Felder, *Tetrahedron Lett.*, 1996, **37**, 1003–1006.

337     V. Krchňák and M. W. Holladay, *Chem. Rev.*, 2002, **102**, 61–92.

338     R. B. Merrifield, *J. Am. Chem. Soc.*, 1963, **85**, 2149–2154.

339     B. Merrifield, *Science (80-. ).*, 1986, **232**, 341.

340     J. C. Nguyen, Y. K. Dzowo, C. Wolfbrandt, J. Townsend, S. Kukatin, H. Wang and M. J. E. Resendiz, *J. Org. Chem.*, 2016, **81**, 8947–8958.

341     M. Adinolfi, G. Barone, L. De Napoli, A. Iadonisi and G. Piccialli, *Tetrahedron Lett.*, 1998, **39**, 1953–1956.

342     R. T. Pon, *Curr. Protoc. Nucleic Acid Chem.*, 2000, **00**, 1–23.

343     G. De Crozals, C. Farre, G. Hantier, D. Léonard, C. A. Marquette, C. A. Mandon, L. Marmuse, C. Louis, J. J. Toulmé, C. Billotey, M. Janier and C. Chaix, *RSC Adv.*, 2012, **2**, 11858–11866.

344     M. Weishaupt, S. Eller and P. H. Seeberger, *Methods Enzymol.*, 2010, **478**, 463–484.

345     R. T. Pon and S. Yu, *Nucleic Acids Res.*, 2004, **32**, 623–31.

346     R. A. Hughes and A. D. Ellington, *Cold Spring Harb. Perspect. Biol.*, 2017, **9**, a023812.

347     C. McCollum and A. Andrus, *Nucleosides and Nucleotides*, 1991, **10**, 573–574.

348     A. Ellington and J. D. Pollard, *Curr. Protoc. Nucleic Acid Chem.*, 2000, **1**, 1.

349     D. A. Wellings, *Methods Enzymol.*, 1997, **289**, 54.

350     R. T. Pon, *Curr. Protoc. Nucleic Acid Chem.*, 2000, **00**, 1–28.

351     TentaGel® M NH₂ Monosized Amino TentaGel Microspheres, http://www.rapp-polymere.com/index.php?id=891.

352     L. K. McKenzie, R. El-Khoury, J. D. Thorpe, M. J. Damha and M. Hollenstein, *Chem. Soc. Rev.*, 2021, **50**, 5126–5164.

353     B. R. Wood, *Chem. Soc. Rev.*, 2016, **45**, 1980–1998.

354     K. Duffy, S. Arangundy-Franklin and P. Holliger, *BMC Biol.*, 2020, **18**, 1–14.

355     D. Voet, J. G. Voet and C. W. Pratt, *Fundam. Biochem.*, 2016, **5**, 36.

356    P. J. Butterworth, *Cell Biochem. Funct.*, 2005, **23**, 293–294.

357    C.-H. Shen, *Diagnostic Mol. Biol.*, 2019, **1**, 1–25.

358    A. Khvorova and J. K. Watts, *Nat. Biotechnol.*, 2017, **35**, 238–248.

359    K. E. Maier and M. Levy, *Mol. Ther. - Methods Clin. Dev.*, 2016, **3**, 16014.

360    J. C. Burnett and J. J. Rossi, *Chem. Biol.*, 2012, **19**, 60–71.

361    C. Rinaldi and M. J. A. Wood, *Nat. Rev. Neurol.*, 2018, **14**, 9–22.

362    R. Rupaimoole and F. J. Slack, *Nat. Rev. Drug Discov.*, 2017, **16**, 203–221.

363    N. Pardi, M. J. Hogan, F. W. Porter and D. Weissman, *Nat. Rev. Drug Discov.*, 2018, **17**, 261–279.

364    W. R. J. D. Galloway, A. Isidro-Llobet and D. R. Spring, *Nat. Commun.*, 2010, 1.

365    L. Gold, B. Polisky, D. Uhlenbeck and M. Yarus, *AnnlL Rev. Biochnn*, 1995, **64**, 763–97.

366    E. Birch-Hirschfeld, R. Egerer, H.-M. Striebel and A. Stelzner, *Collect. Symp. Ser.*, 2002, **5**, 299–303.

367    P. H. Hagedorn, R. Persson, E. D. Funder, N. Albæk, S. L. Diemer, D. J. Hansen, M. R. Møller, N. Papargyri, H. Christiansen, B. R. Hansen, H. F. Hansen, M. A. Jensen and T. Koch, *Drug Discov. Today*, 2018, **23**, 101–114.

368    E. W. M. Ng, D. T. Shima, P. Calias, E. T. Cunningham, D. R. Guyer and A. P. Adamis, *Nat. Rev. Drug Discov.*, 2006, **5**, 123–132.

369    P. Kirsch, *Modern Fluoroorganic Chemistry: Synthesis, Reactivity, Applications*, Wiley Blackwell, 2005.

370    D. M. Lemal, *J. Org. Chem.*, 2004, **69**, 1–11.

371    S. Purser, P. R. Moore, S. Swallow and V. Gouverneur, *Chem. Soc. Rev.*, 2008, **37**, 320–330.

372    C. Heidelberger, N. K. Chaudhuri, P. Danneberg, D. Mooren, L. Griesbach, R. Duschinsky and R. J. Schnitzer, *Nature*, 1957, **179**, 663.

373    S. Purser, P. R. Moore, S. Swallow and V. Gouverneur, 2008, **37**, 1–200.

374    D. L. Roman, C. C. Walline, G. J. Rodriguez and E. L. Barker, *Eur. J. Pharmacol.*, 2003, **479**, 53.

375    J. F. R. Robertson, S. E. Come, S. E. Jones, F. Beex, M. Kaufmann, A. Makris, J. W. R. Nortier, K. Possinger and L.-E. Rutqvist, *Eur. J. Cancer*, 2005, **41**, 346.

376     S. Mabe, J. Eller and W. S. Champney, *Curr. Microbiol.*, 2004, **49**, 248.

377     G. L. Plosker, C. M. Perry and K. L. Goa, *Pharmacoeconomics*, 2001, **19**, 421.

378     G. Dyker, *Angew. Chemie*, 1996, **108**, 2291–2291.

379     G. A. Molander, J. Ham and D. G. Seapy, *Tetrahedron*, 2007, **63**, 768–775.

380     N. Miyaura, T. Ishiyama, H. Sasaki, M. Ishikawa, M. Satoh and A. Suzuki, *J. Am. Chem. Soc.*, 1989, **111**, 314–321.

381     K. C. Nicolaou, P. G. Bulger and D. Sarlah, *Angew. Chemie Int. Ed.*, 2005, **44**, 4442–4489.

382     M. Al-Masum, L. Quinones and L. T. Cain, *Int. J. Org. Chem.*, 2016, **06**, 100–106.

383     N. W. J. Scott, M. J. Ford, C. Schotes, R. R. Parker, A. C. Whitwood and I. J. S. Fairlamb, *Chem. Sci.*, 2019, **10**, 7898–7906.

384     P. Devendar, R.-Y. Qu, W.-M. Kang, B. He and G.-F. Yang, *J. Agric. Food Chem.*, 2018, **66**, 8914–8934.

385     V. S. Sergienko and M. A. Porai-Koshits, *J. Struct. Chem.*, 1988, **28**, 548–552.

386     C. Amatore, A. Jutand and M. A. M'Barki, *Organometallics*, 1992, **11**, 3009–3013.

387     C. Amatore, E. Carré, A. Jutand, M. A. M'Barki and G. Meyer, *Organometallics*, 1995, **14**, 5605–5614.

388     D. Azarian, S. S. Dua, C. Eaborn and D. R. M. Walton, *J. Organomet. Chem.*, 1976, **117**, C55–C57.

389     M. Kosugi, K. Sasazawa, Y. Shimizu and T. Migita, *Chem. Lett.*, 1977, **6**, 301–302.

390     D. Milstein and J. K. Stille, *J. Am. Chem. Soc.*, 1978, **100**, 3636–3638.

391     C. Cordovilla, C. Bartolomé, J. M. Martínez-Ilarduya and P. Espinet, *ACS Catal.*, 2015, **5**, 3040–3053.

392     H. Rosenkrantz, R. W. Fleischman and R. J. Grant, *Toxicol. Appl. Pharmacol.*, 1981, **58**, 118–131.

393     M. L. Crawley and B. M. Trost, *Applications of Transition Metal Catalysis in Drug Discovery and Development*, Wiley, Hoboken, NJ, USA, 2012.

394     J. A. Ragan, J. W. Raggon, P. D. Hill, B. P. Jones, R. E. McDermott, M. J. Munchhof, M. A. Marx, J. M. Casavant, B. A. Cooper, J. L. Doty and Y. Lu, *Org. Process Res. Dev.*, 2003, **7**, 676–683.

395     C. M. So and F. Y. Kwong, *Chem. Soc. Rev.*, 2011, **40**, 4963.

396    A. L. Casado and P. Espinet, *J. Am. Chem. Soc.*, 1998, **120**, 8978–8985.

397    A. L. Casado and P. Espinet, *Organometallics*, 1998, **17**, 954–959.

398    J. F. Hartwig, *Inorg. Chem.*, 2007, **46**, 1936–1947.

399    K. Brunner, J. Harder, T. Halbach, J. Willibald, F. Spada, F. Gnerlich, K. Sparrer, A. Beil, L. Möckl, C. Bräuchle, K.-K. Conzelmann and T. Carell, *Angew. Chemie Int. Ed.*, 2015, **54**, 1946–1949.

400    P. A. Lichtor, Z. Chen, N. H. Elowe, J. C. Chen and D. R. Liu, *Nat. Chem. Biol.*, 2019, **15**, 419–426.

401    G. Vera, B. Diethelm, C. A. Terraza and G. Recabarren-Gajardo, *Molecules*, 2018, **23**, 2051.

402    G. Lu, X. Huangfu, Z. Wu, G. Tang and Y. Zhao, *Adv. Synth. Catal.*, 2019, **361**, 4961–4965.

403    G. Lahoud, V. Timoshchuk, A. Lebedev, K. Arar, Y. M. Hou and H. Gamper, *Nucleic Acids Res.*, 2008, **36**, 6999–7008.

404    E. Ferrer, C. Fàbrega, R. G. Garcia, F. Azorín and R. Eritja, *Nucleosides and Nucleotides*, 1996, **15**, 907–921.

405    T. Onishi, C. Mukai, R. Nakagawa, T. Sekiyama, M. Aoki, K. Suzuki, H. Nakazawa, N. Ono, Y. Ohmura, S. Iwayama, M. Okunishi and T. Tsuji, *J. Med. Chem.*, 2000, **43**, 278–282.

406    W. Kozak, S. Demkowicz, M. Daśko, J. Rachon and J. Rak, *Russ. Chem. Rev.*, 2020, **89**, 281–310.

407    J. R. Mercer, L. H. Xu, E. E. Knaus and L. I. Wiebe, *J. Med. Chem.*, 1989, **32**, 1289–1294.

408    N. Realini, C. Solorzano, C. Pagliuca, D. Pizzirani, A. Armirotti, R. Luciani, M. P. Costi, T. Bandiera and D. Piomelli, *Sci. Rep.*, 2013, **3**, 1035.

409    Y. Uchikubo, T. Hasegawa, S. Mitani, H.-S. Kim and Y. Wataya, *Nucleic Acids Symp. Ser.*, 2002, **2**, 245–246.

410    Y. Yoshida, J. I. Yamamura, H. Sato, M. Koyasu, Y. Obara, H. Sekiguchi, T. Kawana and K. Shiraki, *J. Dermatol. Sci.*, 1996, **13**, 237–241.

411    Y. K. Ho, *Nucleic Acids Res.*, 1984, **12**, 7599–7614.

412    C. Bissantz, B. Kuhn and M. Stahl, *J. Med. Chem.*, 2010, **53**, 5061–5084.

413    N. K. Shinada, A. G. De Brevern and P. Schmidtke, *J. Med. Chem.*, 2019, **62**, 9341–9356.

414    L. M. Salonen, M. Ellermann and F. Diederich, *Angew. Chemie - Int. Ed.*, 2011, **50**, 4808–4842.

415     E. A. Meyer, R. K. Castellano and F. Diederich, *Angew. Chemie Int. Ed.*, 2003, **42**, 1210–1250.

416     R. Fasan, R. L. A. Dias, K. Moehle, O. Zerbe, D. Obrecht, P. R. E. Mittl, M. G. Grütter and J. A. Robinson, *ChemBioChem*, 2006, **7**, 515–526.

417     S. Wang, J. Guo, T. Ono and E. T. Kool, *Angew. Chemie - Int. Ed.*, 2012, **51**, 7176–7180.

418     C. L. Ng, K. Lang, N. A. G. Meenan, A. Sharma, A. C. Kelley, C. Kleanthous and V. Ramakrishnan, *Nat. Struct. Mol. Biol.*, 2010, **17**, 1241–1246.

419     S. El Kazzouli, S. Berteina-Raboin and L. A. Agrofoglio, *Nucleosides, Nucleotides and Nucleic Acids*, 2007, **26**, 1395–1398.

420     C. A. Schalley, in *Analytical Methods in Supramolecular Chemistry*, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2007, pp. 1–16.

421     C. Fonseca Guerra, F. M. Bickelhaupt, J. G. Snijders and E. J. Baerends, *J. Am. Chem. Soc.*, 2000, **122**, 4117–4128.

422     P. Politzer, J. S. Murray and T. Clark, *Phys. Chem. Chem. Phys.*, 2013, **15**, 11178–11189.

423     T. Greene, *Greene's protective groups in organic synthesis*, 2007, vol. 44.

424     L. L. Miller and B. F. Watkins, *J. Am. Chem. Soc.*, 1976, **98**, 1515–1519.

425     M. J. ROBINS, *Tetrahedron Lett.*, 1988, **29**, 2855–2858.

426     C. von Sonntag, *Free-Radical-Induced DNA Damage and Its Repair*, 2006.

427     A. R. Kore, A. Senthilvelan and M. Shanmugasundaram, *Nucleosides, Nucleotides and Nucleic Acids*, 2015, **34**, 92–102.

428     D. Orain, J. Ellard and M. Bradley, *J. Comb. Chem.*, 2002, **4**, 1–16.

429     C. Y. Cho, R. S. Youngquist, S. J. Paikoff, M. H. Beresini, A. R. Hebert, L. T. Berleau, C. W. Liu, D. E. Wemmer, T. Keough and P. G. Schultz, *J. Am. Chem. Soc.*, 1998, **120**, 7706–7718.

430     M. J. Gait and C. Lehmann, *Curr. Protoc. Nucleic Acid Chem.*, 2000, **00**, 1–22.

431     T. Kanamori, A. Takamura, N. Tago, Y. Masaki, A. Ohkubo, M. Sekine and K. Seio, *Org. Biomol. Chem.*, 2017, **15**, 1190–1197.

432    M. Terracciano, I. Rea, L. De Stefano, I. Rendina, G. Oliviero, F. Nici, S. D'Errico, G. Piccialli and N. Borbone, *Nanoscale Res. Lett.*, 2014, **9**, 1–9.

433    M. Hollenstein, *Org. Biomol. Chem.*, 2013, **11**, 5162–5172.

434    F. A. Aldaye and H. F. Sleiman, *Angew. Chemie - Int. Ed.*, 2006, **45**, 2204–2209.

435    A. S. Hansen, A. Thalhammer, A. H. El-Sagheer, T. Brown and C. J. Schofield, *Bioorganic Med. Chem. Lett.*, 2011, **21**, 1181–1184.

436    S. Lingala, L. U. Nordstrøm and P. R. Mallikaratchy, *Tetrahedron Lett.*, 2019, **60**, 211–213.

437    S. Roy and M. Caruthers, *Molecules*, 2013, **18**, 14268–14284.

438    B. N. Gawande, J. C. Rohloff, J. D. Carter, I. Von Carlowitz, C. Zhang, D. J. Schneider and N. Janjic, *Proc. Natl. Acad. Sci. U. S. A.*, 2017, **114**, 2898–2903.

439    K. Sefah, Z. Yang, K. M. Bradley, S. Hoshika, E. Jiménez, L. Zhang, G. Zhu, S. Shanker, F. Yu, D. Turek, W. Tan and S. A. Benner, *Proc. Natl. Acad. Sci. U. S. A.*, 2014, **111**, 1449–1454.

440    D. Proske, M. Blank, R. Buhmann and A. Resch, *Appl. Microbiol. Biotechnol.*, 2005, **69**, 367–374.

441    K. Sefah, D. Shangguan, X. Xiong, M. B. O'Donoghue and W. Tan, *Nat. Protoc.*, 2010, **5**, 1169–1185.

442    F. Tolle and G. Mayer, *Chem. Sci.*, 2013, **4**, 60–67.

443    Z. Yang, D. Hutter, P. Sheng, A. M. Sismour and S. A. Benner, *Nucleic Acids Res.*, 2006, **34**, 6095–6101.

444    T. R. Battersby, D. N. Ang, P. Burgstaller, S. C. Jurczyk, M. T. Bowser, D. D. Buchanan, R. T. Kennedy and S. A. Benner, *J. Am. Chem. Soc.*, 1999, **121**, 9781–9789.

445    L. Gold, D. Ayers, J. Bertino, C. Bock, A. Bock, E. N. Brody, J. Carter, A. B. Dalby, B. E. Eaton, T. Fitzwater, D. Flather, A. Forbes, T. Foreman, C. Fowler, B. Gawande, M. Goss, M. Gunn, S. Gupta, D. Halladay, J. Heil, J. Heilig, B. Hicke, G. Husar, N. Janjic, T. Jarvis, S. Jennings, E. Katilius, T. R. Keeney, N. Kim, T. H. Koch, S. Kraemer, L. Kroiss, N. Le, D. Levine, W. Lindsey, B. Lollo, W. Mayfield, M. Mehan, R. Mehler, S. K. Nelson, M. Nelson, D. Nieuwlandt, M. Nikrad, U. Ochsner, R. M. Ostroff, M. Otis, T. Parker, S. Pietrasiewicz, D. I. Resnicow, J. Rohloff, G. Sanders, S. Sattin, D. Schneider, B. Singer, M. Stanton, A. Sterkel, A. Stewart, S. Stratford, J. D. Vaught, M. Vrkljan, J. J. Walker, M. Watrobka, S. Waugh, A. Weiss, S. K. Wilcox, A. Wolfson, S. K. Wolk, C. Zhang and D. Zichi, *PLoS One*, 2010, **5**, 15004.

446    Á. Furka, F. Sebestyén, M. Asgedom and G. Dibó, *Int. J. Pept. Prot. Res.*, 1991, **37**, 487–493.

447    S. Genheden, A. Reymer, P. Saenz-Méndez and L. A. Eriksson, *Chapter 1. Computational Chemistry and Molecular Modelling Basics*, 2017.

448    A. Pérez, I. Marchán, D. Svozil, J. Sponer, T. E. Cheatham, C. A. Laughton and M. Orozco, *Biophys. J.*, 2007, **92**, 3817–3829.

449    W. Zhang, R. Yang, P. Cieplak, R. Luo, T. Lee, J. Caldwell, J. Wang and P. Kollman, *J. Comput. Chem.*, 2003, **24**, 1999.

450    N. Foloppe and A. D. Mackerell, *J. Comput. Chem.*, 2000, **21**, 86–104.

451    C. Oostenbrink, A. Villa, A. E. Mark and W. F. Van Gunsteren, *J. Comput. Chem.*, 2004, **25**, 1656–1676.

452    S. Myers and A. Baker, *Nat. Biotechnol.*, 2001, **19**, 727–730.

453    A. Manglik, H. Lin, D. K. Aryal, J. D. McCorvy, D. Dengler, G. Corder, A. Levit, R. C. Kling, V. Bernat, H. Hübner, X. P. Huang, M. F. Sassano, P. M. Giguère, S. Löber, D. Duan, G. Scherrer, B. K. Kobilka, P. Gmeiner, B. L. Roth and B. K. Shoichet, *Nature*, 2016, **537**, 185–190.

454    Z. Jin, X. Du, Y. Xu, Y. Deng, M. Liu, Y. Zhao, B. Zhang, X. Li, L. Zhang, C. Peng, Y. Duan, J. Yu, L. Wang, K. Yang, F. Liu, R. Jiang, X. Yang, T. You, X. Liu, X. Yang, F. Bai, H. Liu, X. Liu, L. W. Guddat, W. Xu, G. Xiao, C. Qin, Z. Shi, H. Jiang, Z. Rao and H. Yang, *Nature*, 2020, **582**, 289–293.

455    X. Hu, J. H. Shrimp, H. Guo, M. Xu, C. Z. Chen, W. Zhu, A. V. Zakharov, S. Jain, P. Shinn, A. Simeonov, M. D. Hall and M. Shen, *ACS Pharmacol. Transl. Sci.*, 2021, **4**, 1124–1135.

456    E. Glaab, *Brief. Bioinform.*, 2016, **17**, 352–366.

457    T. I. Adelusi, A. Q. K. Oyedele, I. D. Boyenle, A. T. Ogunlana, R. O. Adeyemi, C. D. Ukachi, M. O. Idris, O. T. Olaoba, I. O. Adedotun, O. E. Kolawole, Y. Xiaoxing and M. Abdul-Hammed, *Informatics Med. Unlocked*, 2022, **29**, 100880.

458    H. van de Waterbeemd and E. Gifford, *Nat. Rev. Drug Discov.*, 2003, **2**, 192–204.

459    M. J. Hansen, W. A. Velema, M. M. Lerch, W. Szymanski and B. L. Feringa, *Chem. Soc. Rev.*, 2015, **44**, 3358–3377.

460    A. Rodrigues-Correia, X. M. M. Weyel and A. Heckel, *Org. Lett.*, 2013, **15**, 5500–5503.

461    G. Zhou, F. Khan, Q. Dai, J. E. Sylvester and S. J. Kron, *Mol. Biosyst.*, 2012, **8**, 2395–2404.

462    S. V. Wegner, O. I. Sentürk and J. P. Spatz, *Sci. Rep.*, 2015, **5**, 1–7.

463    T. Kakiyama, K. Usui, K. Y. Tomizaki, M. Mie, E. Kobatake and H. Mihara, *Polym. J.*, 2013, **45**, 535–539.

464    R. MacArron, M. N. Banks, D. Bojanic, D. J. Burns, D. A. Cirovic, T. Garyantes, D. V. S. Green, R. P. Hertzberg, W. P. Janzen, J. W. Paslay, U. Schopfer and G. S. Sittampalam, *Nat. Rev. Drug Discov.*, 2011, **10**, 188–195.

465    M. S. Butler, *J. Nat. Prod.*, 2004, **67**, 2141–2153.

466    D. J. Newman and G. M. Cragg, *J. Nat. Prod.*, 2007, **70**, 461–477.

467    R. B. Merrifield, *Excerpta Med., I.C.S.*, 1976, **374**, 29–39.

468    L. M. Mayr and D. Bojanic, *Curr. Opin. Pharmacol.*, 2009, **9**, 580–588.

469    J. Inglese, R. L. Johnson, A. Simeonov, M. Xia, W. Zheng, C. P. Austin and D. S. Auld, *Nat. Chem. Biol.*, 2007, **3**, 466–479.

470    R. Kramer and D. Cohen, *Nat. Rev. Drug Discov.*, 2004, **3**, 965–972.

471    M. S. Attene-Ramos, C. P. Austin and M. Xia, *High Throughput Screening*, Elsevier, Third Edit., 2014, vol. 2.

472    G. Pegoraro and T. Misteli, *Trends Genet.*, 2017, **33**, 604–615.

473    D. P. O'Halloran, K. Wynne and J. A. Geoghegan, *Infect. Immun.*, 2015, **83**, 1598–1609.

474    S. A. Moon, M. K. Ki, S. Lee, M. L. Hong, M. Kim, S. Kim, J. Chung, S. G. Rhee and H. Shim, *Mol. Cells*, 2011, **31**, 509–513.

475    V. Peters and B. H. A. Rehm, *J. Biotechnol.*, 2008, **134**, 266–274.

476    M. Mishra, S. Tiwari, A. Gunaseelan, D. Li, B. D. Hammock and A. V. Gomes, *Electrophoresis*, 2019, **40**, 1731–1739.

477    T. Lakshmipriya, S. C. B. Gopinath and T. H. Tang, *PLoS One*, 2016, **11**, 16–20.

478    D. Xu and S. V. Wegner, *Chem. Sci.*, 2020, **11**, 4422–4429.

479    X. Y. Lu, D. C. Wu, Z. J. Li and G. Q. Chen, in *Progress in Molecular Biology and Translational Science*, Elsevier B.V., 2011, vol. 104, pp. 299–323.

480     H. Berman, K. Henrick and H. Nakamura, *Nat. Struct. Mol. Biol.*, 2003, **10**, 980–980.

481     M. Zondlo Fiume, *Int. J. Toxicol.*, 2001, **20**, 21–45.

482     O. H. Laitinen, V. P. Hytönen, H. R. Nordlund and M. S. Kulomaa, *Cell. Mol. Life Sci.*, 2006, **63**, 2992–3017.

483     K. J. Roux, D. I. Kim, M. Raida and B. Burke, *J. Cell Biol.*, 2012, **196**, 801–810.

484     P. C. Weber, D. H. Ohlendorf, J. J. Wendoloski and F. R. Salemme, *Science (80-. ).*, 1989, **243**, 85–88.

485     M. Murphy, Kenneth. Travers, Paul. Walport, *Immunobiology*, 7th edn., 2008.

486     J. F. J. Todd, *Int. J. Mass Spectrom. Ion Process.*, 1995, **142**, 209–240.

487     P. Thakur, S. Abernathy, P. A. Limbach and B. Addepalli, *Locating chemical modifications in RNA sequences through ribonucleases and LC-MS based analysis*, Elsevier Inc., 1st edn., 2021, vol. 658.

488     S. C. Pomerantz and J. A. McCloskey, *Methods Enzymol.*, 1990, **193**, 796–824.

489     J. A. Kowalak, S. C. Pomerantz, P. F. Crain and J. A. Mccloskey, *Nucleic Acids Res.*, 1993, **21**, 4577–4585.

490     A. Björkbom, V. S. Lelyveld, S. Zhang, W. Zhang, C. P. Tam, J. C. Blain and J. W. Szostak, *J. Am. Chem. Soc.*, 2015, **137**, 14430–14438.

491     G. Schutz, S. Kieval, B. Groner, A. E. Sippel, D. T. Kurtz and P. Feigelson, *Cellulose*, 1977, **4**, 71–84.

492     J. Shendure and H. Ji, *Nat. Biotechnol.*, 2008, **26**, 1135–1145.

493     P. J. Sample, K. W. Gaston, J. D. Alfonzo and P. A. Limbach, *Nucleic Acids Res.*, 2015, **43**, 1–13.

494     A. Nyakas, S. R. Stucki and S. Schürch, *J. Am. Soc. Mass Spectrom.*, 2011, **22**, 875–887.

495     T. Hiroshi Nakayama, Nobuhiro Takahashi and T. Isobe, *Mass Spectrom. Rev.*, 2011, **30**, 1000–1012.

496     S. C. Pomerantz, J. A. Kowalak and J. A. McCloskey, *J. Am. Soc. Mass Spectrom.*, 1993, **4**, 204–209.

497     J. Rozenski and J. A. McCloskey, *J. Am. Soc. Mass Spectrom.*, 2002, **13**, 200–203.

498     A. Nyakas, L. C. Blum, S. R. Stucki, J. L. Reymond and S. Schürch, *J. Am. Soc. Mass Spectrom.*, 2013, **24**, 249–256.

499     R. Matthiesen and F. Kirpekar, *Nucleic Acids Res.*, 2009, **37**, e48–e48.

500     H. Nakayama, M. Akiyama, M. Taoka, Y. Yamauchi, Y. Nobe, H. Ishikawa, N. Takahashi and T. Isobe, *Nucleic*

*Acids Res.*, 2009, **37**, e47–e47.

501     E. A. Kapp, F. Schütz, L. M. Connolly, J. A. Chakel, J. E. Meza, C. A. Miller, D. Fenyo, J. K. Eng, J. N. Adkins, G. S. Omenn and R. J. Simpson, *Proteomics*, 2005, **5**, 3475–3490.

502     H. Oberacher and F. Pitterl, *Int. J. Mass Spectrom.*, 2011, **304**, 124–129.

503     S. et Al., *J Mass Spectrom*, 2012, **78**, 711–716.

504     M. Kretschmer, G. Lavine, J. McArdle, S. Kuchimanchi, V. Murugaiah and M. Manoharan, *Anal. Biochem.*, 2010, **405**, 213–223.

505     M. Jora, P. A. Lobue, R. L. Ross, B. Williams and B. Addepalli, *Biochim. Biophys. Acta - Gene Regul. Mech.*, 2019, **1862**, 280–290.

506     Binding affinity, https://www.malvernpanalytical.com/en/products/measurement-type/binding-affinity.

507     T. D. Pollard, *Mol. Biol. Cell*, 2010, **21**, 4061–4067.

508     D. C. Swinney, *Molecular Mechanism of Action (MMoA) in Drug Discovery*, Elsevier Inc., 1st edn., 2011, vol. 46.

509     A. Tiwari and S. J. Dhoble, *Talanta*, 2018, **180**, 1–11.

510     X. Hun, B. Liu and Y. Meng, *Microchim. Acta*, 2017, **184**, 3953–3959.

511     R. F. Delgadillo, T. C. Mueser, K. Zaleta-Rivera, K. A. Carnes, J. González-Valdez and L. J. Parkhurst, *Detailed characterization of the solution kinetics and thermodynamics of biotin, biocytin and HABA binding to avidin and streptavidin*, 2019, vol. 14.

512     J. B. Pawley, *Biological Confocal Microscopy*, 2006.

513     G. Kahl, in *The Dictionary of Genomics, Transcriptomics and Proteomics*, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2015, p. 1.

514     H. Schneckenburger, *Methods Appl. Fluoresc.*, 2019, **8**, 013001.

515     B. T. Dye, *Clin. Appl. Immunol. Rev.*, 2005, **5**, 307–324.

516     R. H. Batchelor, A. Sarkez, W. G. Cox and I. Johnson, *Biotechniques*, 2007, **43**, 503–507.

517     M. Zaldini Hernandes, S. T. Melo Cavalcanti, D. M. Rodrigo Moreira, W. Filgueira de Azevedo Junior and A.

Cristina Lima Leite, *Curr. Drug Targets*, 2010, **11**, 303–314.

518     J. Byun, *Life*, 2021, **11**, 1–18.

519     S. Qian, D. Chang, S. He and Y. Li, *Anal. Chim. Acta*, 2022, **1196**, 339511.

520     Aptamer Market, https://www.marketdataforecast.com/market-reports/aptamers-market, (accessed 20 June 2022).

521     J. Zhou, J. Rossi, C. Biology and E. D. Rd, *Nat Rev Drug Discov*, 2017, **8**, 181–202.

522     S. Yoon and J. J. Rossi, *Expert Opin. Drug Discov.*, 2017, **12**, 317–319.

523     Aptamer Group and AstraZeneca, https://aptamergroup.com/aptamer-group-announces-agreement-with-astrazeneca-to-explore-next-generation-drug-delivery-devices/, (accessed 20 June 2022).

524     D. Buffery, *Am. Heal. Drug Benefits*, 2015, **8**, 216–222.

525     Y. Morita, M. Leslie, H. Kameyama, D. E. Volk and T. Tanaka, *Cancers (Basel).*, , DOI:10.3390/cancers10030080.

526     R. A. de Mello, N. M. Neves, H. Tadokoro, G. A. Amaral, P. Castelo-Branco and V. A. de A. Zia, *J. Clin. Med.*, 2020, **9**, 1–20.

527     B. Melosky and V. Hirsh, *Front. Oncol.*, 2014, **4 AUG**, 1–6.

528     S. Wang, B. Mao, M. Wu, J. Liang and L. Deng, *Antonie van Leeuwenhoek, Int. J. Gen. Mol. Microbiol.*, 2018, **111**, 199–208.

529     C. A. Stein and D. Castanotto, *Mol. Ther.*, 2017, **25**, 1069–1075.