Novel Methods for Probing and Detecting Nucleic Acids

by

Jake Giovanni Carter

A thesis submitted to the

University of Birmingham

For the degree of

Doctor of Philosophy

School of Chemistry
College of Engineering and Physical Sciences
University of Birmingham
January 2022

Abstract

Nucleic acids such as Deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) are essential for life in all organisms. As such the detection of nucleic acids can prove to be a useful tool in identifying the cause of an illness. The current method for detecting the presence of small quantities of nucleic acids is hinged on the polymerase chain reaction (PCR) the 'gold standard' of amplification, however, newer isothermal techniques that do not rely on temperature cycling such as loop mediated isothermal amplification (LAMP) are becoming more popular. However, the detection of the small quantities of nucleic acids found in pathogens still requires a lengthy process to increase the quantity of material prior to detection. Novel methods in which to amplify and subsequently detect nucleic acids were explored over the course of this research. Through the combination of a rapid isothermal amplification technique, and an established method that uses linear dichroism (LD) as a readout we demonstrated the ability to detect 10 fM of DNA in under 15 minutes. The detection of RNA requires a reverse transcription step to convert the RNA into DNA in order to be detected, we were able to develop an alternative method to detect the presence of viral RNA without the requirement of a length RT step. The results of this work demonstrate how the application of a reverse transcription free (RTF) step prior to an isothermal amplification technique, enables the detection of pathogenic material in less than 10 minutes. This offers an alternative, simpler, method for point of care (POC) testing. Work within the group has previously demonstrated the ability to detect DNA using M13 Bacteriophage based biosensors that are monitored using LD, unfortunately the ability to scale up production is limited; therefore, we investigated the ability of a synthetic foldamers to be measured using LD, offering a both a novel method for probing not only the structure of the foldamer itself, but also as a possible method for detecting nucleic acids.

It’s just a book. No harm ever came from reading a book.

Acknowledgements

This is probably the hardest part of the entire thesis for me to write, as I know there are so many people I would like to thank, yet I don't feel I am able to find the appropriate words to express how grateful I am to each and every one, but I suppose I should at least try.

I would like to start by thanking Nicola, you have been there with me throughout this entire journey, I couldn't have done this without your support , and I am so grateful to have you in my life and I hope to spend it with you. Thank you to my mum and dad, Helen, and Ian for being there, supporting me throughout all my endeavours and being interested in what I have been doing, I would also like to thank Jane, Bob, and Darren for being a huge part of my life and showing such an interest. I would also like to show my gratitude for my grandad Joe and my nan Jean for their love and support in all aspects of my life.

I would whole heartly like to thank Professor James Tucker and Professor Timothy Dafforn, for enabling me to undertake this PhD you have both been an amazing source of inspiration and your constant belief in my abilities as a scientist are how I was able to get here; working with you both has been my absolute pleasure, thank you both for being the best supervisors I could have asked for. To all the Tucker and Dafforn group, past and present; Charlotte, Jack, Francia, Ed, Georgina, Marium, YiFeng, Liyao, Aldrich, Klaudia, Aysha, Lydia, Eva, La, Zoe, Matt, James, Charles, Steph, Khalid, Bethan, Sarah, Naomi, and Liam thank you all for being a source of laughter, inspiration and proving a welcome break. However, it was not just members of the group who have been invaluable Jean-Louis, Bert, Rueben, Alison, Matt, Lorea, Andrew, Celina and Charlie I have had the privilege of working alongside you at one time or another, thank you for your support and amazing scientific abilities. If the opportunity arises, I look forward to working with you in the future. I would like to express my thanks to the all the staff in the school of chemistry and biosciences that have helped me along the way.

Nicola, Hazel, and Danny, I have known you all since before any of us started at Birmingham and I am glad we were able to progress through our PhDs at the same time. I wish you all the best of luck with your theses. Trevor, Lizzie, Jamie, Nick, Menisha, Richard, Conner, Izzy, Alex, Harry, George, Holly, and Sam while I don't know half of you half as well as I should like, I am lucky to be able to call you all friends. There are countless more people that I am sure I have forgotten, and I am sorry if you are reading this and do not see your name, it does not mean you are any less important.

The past four years have gone by so quickly and while they haven't been without sadness and a pandemic, I can say that the majority of this time has been filled with moments I hope I never forget; and on that note let us begin.

Table of Contents

Abbreviations

$A_{\parallel}$	Absorbance of Parallel Light
$A_{\perp}$	Absorbance of Perpendicular Light
A_{iso}	Absorbance of Isotropic Light
A_l	Absorbance of Left-Handed Light
A_r	Absorbance of Right-Handed Light
AC	Alternating Current
ACE2	Angiotensin-Converting Enzyme 2
ATP	Adenosine Triphosphate
bp	Base Pairs
CD	Circular Dichroism
cDNA	Complementary Deoxyribonucleic Acid
COVID-19	Coronavirus Disease 2019
CuAAC	Copper Catalysed Huisgen Cycloaddition Reaction
CV	Cyclic Voltammetry
DCM	Dichloromethane
dCTP	Deoxycytidine Triphosphate
dCTPαS	Deoxycytidine Thiotriphosphate
ddPCR	Digital Droplet Polymerase Chain Reaction
DIBMA	Poly(Diisobutylene-*alt*-Maleic acid)
DMF	*N, N*-Dimethylformamide
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic Acid
dNTP	Deoxynucleoside Triphosphate
dsDNA	Double Stranded Deoxyribonucleic Acid

E. coli	*Escherichia coli*
ESI	Electrospray Ionisation
ESI-MS	Electrospray Ionisation Mass Spectrometry
EXPAR	Exponential Amplification Reaction
FDR	False Discovery Rate
FN	False Negative
FP	False Positive
GCPR	G-Coupled Protein Receptors
hCG	Human Chorionic Gonadotropin
HDA	Helicase Dependant Amplification
HOMO	Highest Occupied Molecular Orbital
HPLC	High Performance Liquid Chromatography
HPV	Human Papillomaviruses
HRMS	High Resolution Mass Spectrometry
IC	Internal Conversion
IR	Infrared
ISC	Intersystem Crossing
LAMP	Loop-Mediated Isothermal Amplification
LD	Linear Dichroism
LD^r	Reduced Linear Dichroism
LFT	Lateral Flow Tests
LUMO	Lowest Unoccupied Molecular Orbital
MB	Molecular Beacon
MERS	Middle Eastern Respiratory Syndrome
MHRA	Medicines and Healthcare Products Regulatory Agency

miRNA	Micro Ribonucleic Acid
m.p.	Melting Point
mRNA	Messenger Ribonucleic Acid
MS	Mass Spectrometry
NAAT	Nucleic Acid Amplification Test
ncRNA	Non-Coding Ribonucleic Acid
NHS	National Health Service
NHS	*N*-Hydroxysuccinimide
NMR	Nuclear Magnetic Resonance
PCR	Polymerase Chain Reaction
PE	Polyethylene
PE^{OX}	Oxidised Polyethylene
PEG	Polyethylene Glycol
PEM	Photoelastic Modulator
PHE	Public Health England
P(*p*-AT)	Poly(*para*-Aryltriazole)
PPV	Positive Predictive Value
qPCR	Quantitative Polymerase Chain Reaction
R_0	Basic Reproductive Number
RBD	Receptor-Binding Domain
RCA	Rolling Circle Amplification
RNase MRP	Ribonuclease for Mitochondrial Ribonucleic Acid Processing
RNA	Ribonucleic Acid
RP-HPLC	Reversed-Phase High Performance Liquid Chromatography
rRNA	Ribosomal Ribonucleic Acid

rt	Room Temperature
RT	Reverse Transcriptase
RT-LAMP	Reverse Transcription Loop-Mediated Isothermal Amplification
RT-PCR	Reverse Transcription Polymerase Chain Reaction
RTF-EXPAR	Reverse Transcription Free Exponential Amplification Reaction
SARS	Severe Acute Respiratory Syndrome
SARS-CoV-2	Severe Acute Respiratory Syndrome Coronavirus 2
SDA	Strand Displacement Amplification
siRNA	Small Interfacing Ribonucleic Acid
SMA	Poly(Styrene-*co*-Maleic Acid)
SMCC	Succinimidyl 4-(*N*-maleimidomethyl) cyclohexane-1-carboxylate
SMI	Poly(Styrene-*co*-(*N*-(3-*N*′,*N*′-Dimethylaminopropyl)Maleimide))
SPR	Surface Plasmon Resonance
SSB	Single Stranded Binding Protein
ssDNA	Single Stranded Deoxyribonucleic Acid
TEM	Transmission Electron Microscopy
T_m	Melting Temperature
TMV	Tobacco Mosaic Virus
TN	True Negative
TP	True Positive
tRNA	Transfer Ribonucleic Acid
UV-Vis	Ultraviolet Visible
WHO	World Health Organisation
wt	Wild Type
λ_{max}	Lambda Max

μ	Electron Transition Dipole Moment
σ	Standard Deviation

Nucleic acid bases:

A	Adenine
C	Cytosine
G	Guanine
T	Thymine
U	Uracil

Chapter 1: Introduction

1.1 Introduction

This chapter serves as an introduction to the detection of biomarkers, with particular interest paid to nucleic acids, their structure, and the major detection methods. Finally, there will be an introduction to methods of amplification, an area of focus within this thesis.

1.2 Detection of Biomarkers

Infection and disease are detrimental to the health of the afflicted individual. A faster diagnosis of the cause of illness can lead to faster treatment and a higher chance of survival.[1] During an illness a range of biological molecules, called biomarkers, produced either by the host as an immune response or by the cause of the illness can be used to indicate the cause.[2] These biomarkers can range from proteins on the surface of a cells, antibodies produced by the host, to the genetic material of the cells; all of which can be used to diagnose the cause of the sickness. As a result, research into diagnostics is an ever growing, multi-billion dollar industry.

1.3 Protein Detection

Protein detection provides a simple method of identifying the cause of illness. While there a plethora of methods to achieve this goal; currently one of the most popular method for the detection of proteins, is using lateral flow antigen tests.

1.3.1 Lateral Flow Assay

Lateral flow tests (LFT), are a method of detection that relies on the binding affinity of antigen – antibody complexes to rapidly detect the presence of a range of analytes.[3,4] LFTs use capillary action to move a sample along a strip of nitrocellulose and if the target is present, it binds to a test line giving a visible readout.[5–8] Briefly explained, (**Figure 1.1**) a liquid sample is placed on the sample pad, this disperses the samples evenly whilst filtering off unwanted elements, such as blood cells and food.[3,9–11] The capillary action forces the sample along the strip to the next station, the conjugation pad. The conjugation pad contains monoclonal antibodies, specific to the target, that are tagged with a label (fluorophores,[10,12] nanoparticles[9,13] etc).[6,14–17] If the target analyte is present it binds to a labelled antibody forming an analyte – antibody complex, this is then carried to the membrane. The membrane contains at least two regions, the test line, and the control line. The first region, the test line, contains a secondary antibody that interacts with the analyte of the analyte – antibody complex, halting its progress along the membrane. The accumulation of labelled complexes in this region gives rise to a coloured line indicating a positive test.[18–20] Meanwhile, the control line contains a recognition element for the labelled antibody itself. This gives rise to another coloured line which is used to ensure the LFT is functioning in the intended manner.[6,8,10] At the terminal end of the strip there is an absorbent pad, to remove excess reagents and to prevent back flowing of the liquid.[9,15,17]

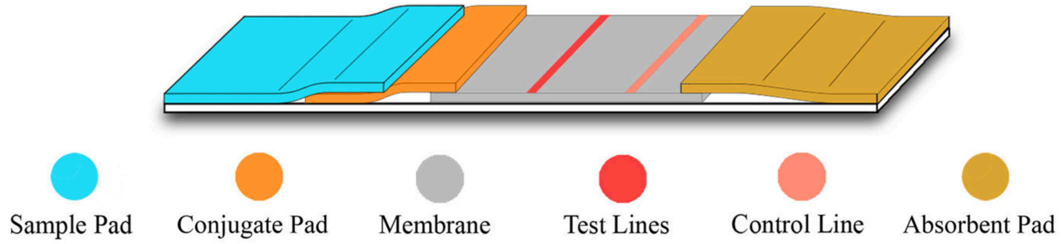


Figure 1.1: *Diagram representing the various components of a lateral flow test. Image taken from Chen et al* [21]

The simplicity of the system has enabled the detection of a plethora of analytes ranging from hormones, and markers of infection to more exotic analytes such pathogens themselves or markers for cardiovascular diseases[3,22] and cancer.[23] Until recently, the most common LFT was a pregnancy test, which detects the hormone human chorionic gonadotropin (hCG) that is produced in elevated levels during early gestation (10 to 25 mIU/mL). As hCG is eliminated *via* urine, a pregnancy test has been tailored to detect this hormone. As with other lateral flow assays a tagged antibody binds to the target, flows down the strip, and binds to the test line creating a visible result, thus confirming pregnancy.[24] Interestingly during the Ebola epidemic of 2014 – 2015 a LFT was developed that was capable of detecting the Ebola antigen ReEBOV VP40.[25] Prior to the introduction of this LFT, diagnosis was based solely on reverse transcription polymerase chain reaction (RT-PCR), the LFT was recognised as a rapid and accurate diagnostic platform that was critical to containing and eliminating the Ebola outbreak.[26]

1.4 Nucleic Acid Detection

While protein detection offers specific antigen – antibody binding detection one downfall of protein detection is if the protein is at low levels it is difficult to accurately identify the presence of a target analyte. Another molecule of interest for detection are nucleic acids, Deoxyribonucleic acid (DNA) and Ribonucleic acid (RNA) are the genetic material for the living world as we know it.[27] These molecules offer high accuracy and can be designed specifically for each individual disease or infection.

1.4.1 Structure of Deoxyribonucleic Acid (DNA)

Nearly every cell contains DNA,[28] from simple single celled organisms to complex multicellular lifeforms.[29] The information on which life is based is stored within DNA as a biological code, constructed from four nitrogen-containing chemical bases. These consist of two purine based groups: Adenine (A) and Guanine (G), and two pyrimidines based groups: Cytosine (C) and Thymine (T) (**Figure 1.2**).[30]

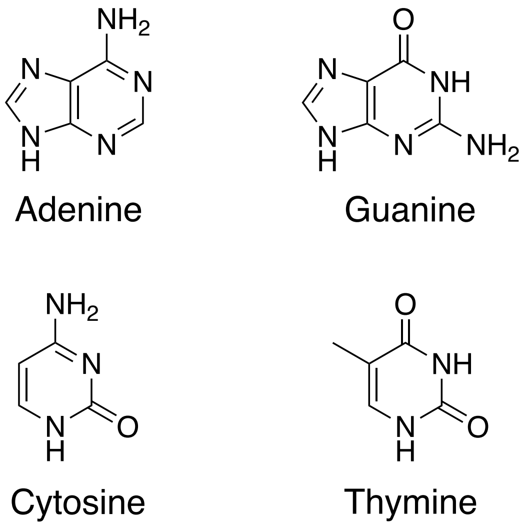


***Figure 1.2**: Nitrogenous bases found in nucleic acids. Top row displays the purine based group (Left) Adenine (Right) Guanine. Bottom row displays the pyrimidine based groups (Left) Cytosine (Right) Thymine.*

The order in which these bases appear determines the structure produced and the function they carry out.[29,30] DNA itself is constructed from repeating deoxyribonucleotide units (**Figure 1.3**); a base is attached to a five membered deoxyribose sugar ring to which a phosphate group is attached.[31]

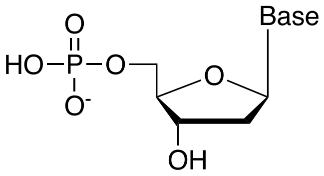


***Figure 1.3**: The general deoxyribonucleotide structure*

Deoxyribonucleotides possess a free phosphate group at the 5´ position and a free hydroxyl group at the 3´.[32] Through a covalent interaction between these two group multiple deoxyribonucleotides are linked together forming a chain which make the ‘backbone’ of the polymer and thus a single strand of DNA (ssDNA).[27,29,30] The interaction between phosphate and hydroxyl groups are not the only interactions that occur; when two ssDNA oligomers come together they are able to pair with other bases through hydrogen bonding. A and T link through 2 hydrogen bonds while C and G link through 3 hydrogen bonds (**Figure 1.4**).

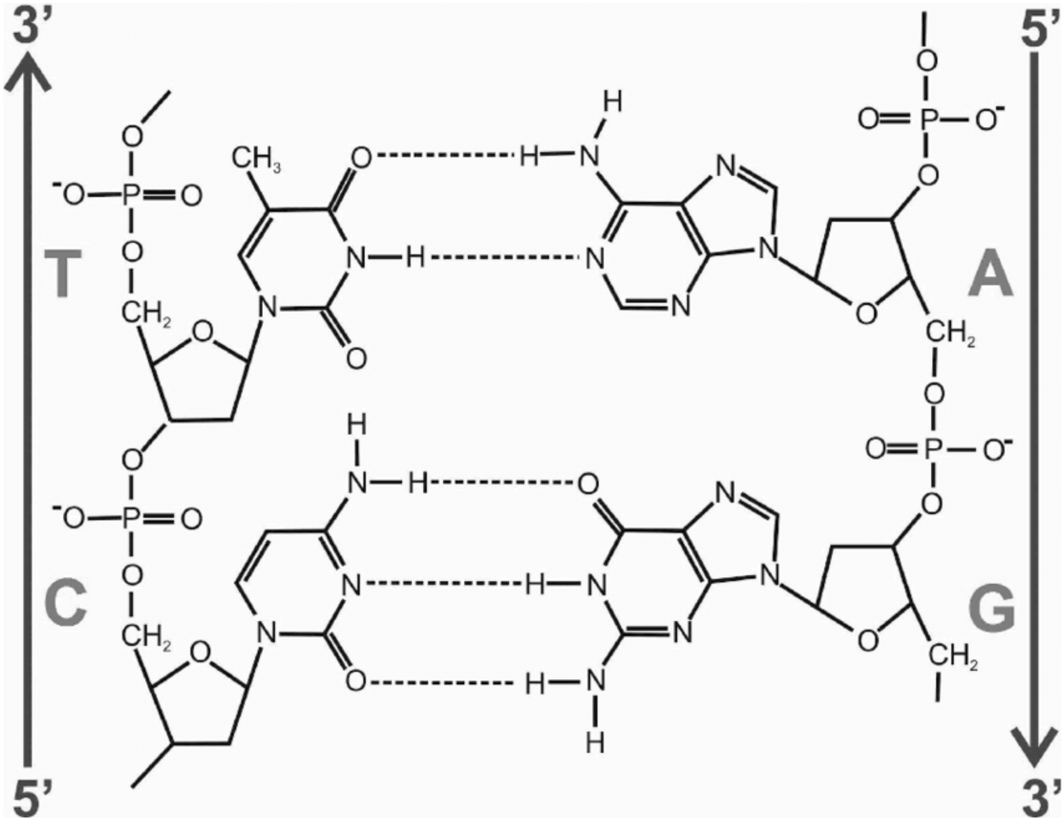


Figure 1.4*: DNA structure showing the hydrogen bonds between the complementary bases two bonds between T and A and three bonds between C and G. Image taken from Harding et al.*[33]

The hydrogen bonding between the bases causes the formation of double-stranded DNA (dsDNA).[27,30–32] The base pairing between complimentary bases enables the DNA to pack into the most energetically stable conformation, with the bases on the interior of the dsDNA ensuring the sugar-phosphate backbone is evenly spaced along the length of the DNA molecule.[30,32] To further maximise base stacking efficiency, the two sugar-phosphate backbones twist around each other to form an antiparallel double helix structure, with a complete turn occurring every ten base pairs (bp).[31,34] This double helix forms the well-known structure of DNA.

1.4.2 Structure of Ribonucleic Acid (RNA)

DNA is not the only nucleic acid; there is an additional analogue species called RNA, which plays just as fundamental of a part in the synthesis of life. While DNA contains the code of life, RNA enables this code to be used within a cell by acting as a messenger delivering the blueprints to the 'factory', the ribosome.[35] RNA is fundamentally the same as DNA in regard to structure, albeit with two major differences (**Figure 1.5**).

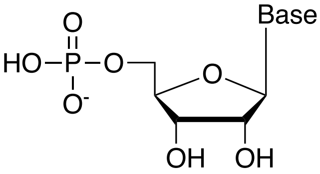
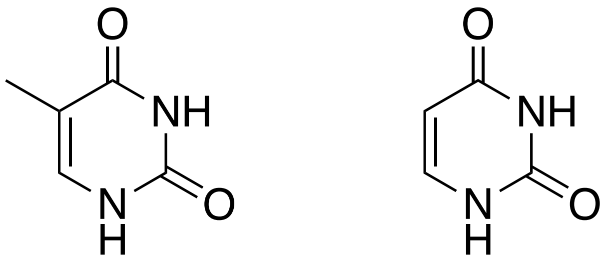


Figure 1.5: *The general ribonucleotide structure*

Both DNA and RNA are composed of a sugar-phosphate backbone with a nitrogen-containing base linked to the sugar ring.[36] Where RNA differs is firstly seen in the sugar ring. DNA possesses a five-membered deoxyribose sugar ring, while RNA possesses a five-membered ribose sugar ring which contains an additional hydroxy group at the 2´ position.[37] This added hydroxy group reduces the stability of the RNA by increasing the likelihood of hydrolysis *in vivo*.[36,38] The hydroxy group at the 2´ position is deprotonated, where it is able to attack the adjacent phosphorus and thus break the phosphodiester bond, causing the RNA backbone to be cleaved.[31,36,38,39] The second major difference is to the nitrogen-containing bases. While A, C and G are found within both DNA and RNA, T is substituted for Uracil (U) in RNA. Structurally U is similar to T, except it is lacking a methyl group (**Figure 1.6**).



Thymine Uracil

__Figure 1.6__: Structure of the nucleic acid bases Thymine and Uracil.

During DNA synthesis U bases are methylated to form T bases during incorporation.[40] While it is still unclear why these bases differ between DNA and RNA, it is believed that the additional methyl group increases the stability of DNA. Many organisms have evolved to product a Uracil – DNA glycosylase enzyme which selectively removes U from a dsDNA sequence. If U was incorporated naturally in the correct place, it would be removed during DNA synthesis.[41] In terms of structure RNA varies greatly from DNA. DNA is naturally composed of two strands creating a duplex, while on the other hand RNA occurs as a single strand and folds in on itself, forming complex structures comparable to proteins (**Figure 1.7**).[42] While DNA may be burdened with glorious purpose, alone it is unable to complete what destiny intended. The DNA must first be converted into RNA *via* a process called transcription. While around 80% of the human genome can be converted into RNA, only ~2% is converted into the coding messenger RNA (mRNA) which is able to produce proteins. The remaining ~78% creates non-coding RNA (ncRNA) which is incapable of coding for proteins.[43] There are a plethora of ncRNA species, which serve a range of functions such as gene regulation, microRNA (miRNA)[44], small interfacing RNA (siRNA)[45], and DNA replication with ribonuclease for mitochondrial RNA processing (RNase MRP)[46]. Some of these ncRNA species are fundamental for converting mRNA into proteins, and these include transfer RNA (tRNA) and ribosomal RNA (rRNA) (**Figure 1.7**), all of which offer a potential target for nucleic acid detection.

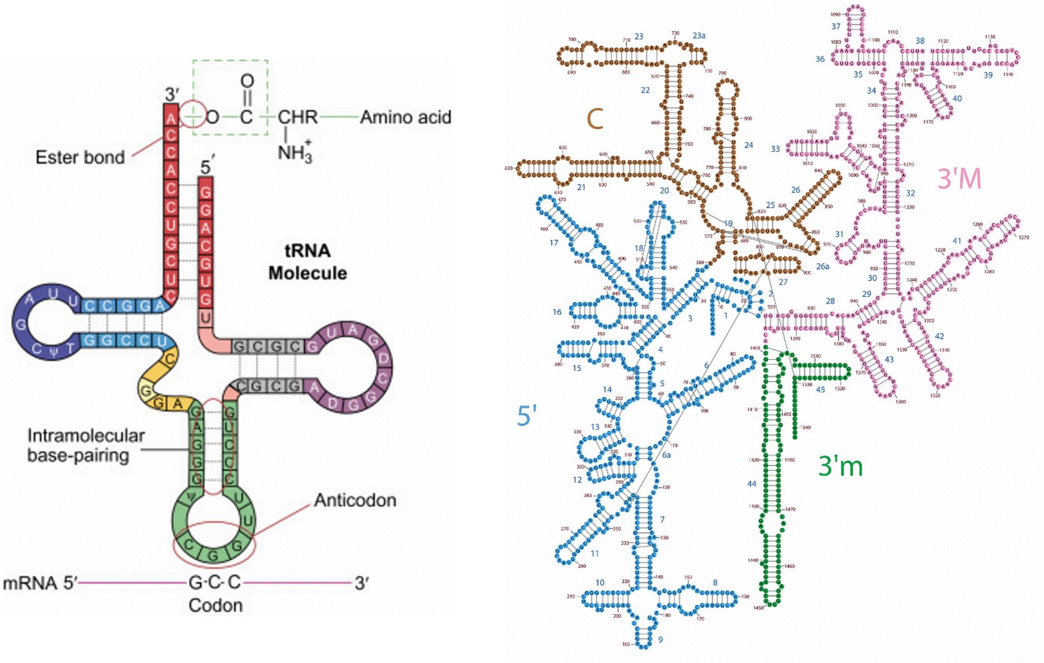


Figure 1.7: *(Left) The secondary structure of tRNA, image adapted from Litwack.*[47] *(Right) The secondary structure of rRNA, image adapted from Petrov et al.*[48]

1.4.3 Nucleic Acid detection

The detection of nucleic acids, be that DNA or RNA, is an efficient and accurate way to detect the presence of a pathogen or a genetic disease. There are a range of methods which can be employed, below some of the most popular methods will be discussed.

1.4.3.1 Electrochemical Detection

Electrochemical sensing monitors the change in potential of molecules which are able to be reduced or oxidised when a voltage is applied.[49] There are numerous ways in which this change can be monitored, including linear sweep voltammetry[50] and square wave voltammetry,[51] however the most common method is cyclic voltammetry (CV).[49] CV employs the use of three electrodes in a single cell: the working electrode, the counter electrode, and the reference electrode. The working electrode is where the redox reaction takes place. A current passes to or from this electrode (depending on the nature of the analyte) and the change in current is recorded by the potentiostat. This electrode is usually made from either gold,[52] boron doped diamond[53], or glassy carbon.[54] In the case of nucleic acid sensing the electrode is usually gold with a thiolated ssDNA strand conjugated to the gold surface.[55] The counter electrode completes the circuit and, depending on the nature of the analyte, acts as either an electron source or electron sink for the analyte to be reduced or oxidised; therefore, the current at this electrode is not measured. For every redox reaction an electron is transferred between the working electrode and counter electrode in order to maintain a charge balance, and because of the requirement for fast electron transfer the counter electrode is usually composed of platinum.[49] The final electrode is the reference electrode, and this along with the working electrode is monitored by the potentiostat, however for this electrode the voltage is not varied. As the name suggests, the purpose of the reference electrode is to act as a reference for the working electrode. Therefore, the reference electrode is usually composed of Ag/AgCl, (silver wire in a solution of silver chloride), this choice meaning the potential remains constant for the duration of the experiment.[49,52,54] All three electrodes are immersed in an electrolyte solution and as mentioned, the change of potential is measured at the working electrode *via* a potentiostat. By monitoring the change in potential, CV enables information about redox activity of a molecule to be gained. The cycling of the voltage will lead to the production of a voltammogram (**Figure 1.8**) which possess a peak as a result of oxidation (positive peak) or reduction (negative peak); if the analyte can be reverted to its original redox state a secondary peak will also be produced on the opposite side as seen in **Figure 1.8**.[49,56]

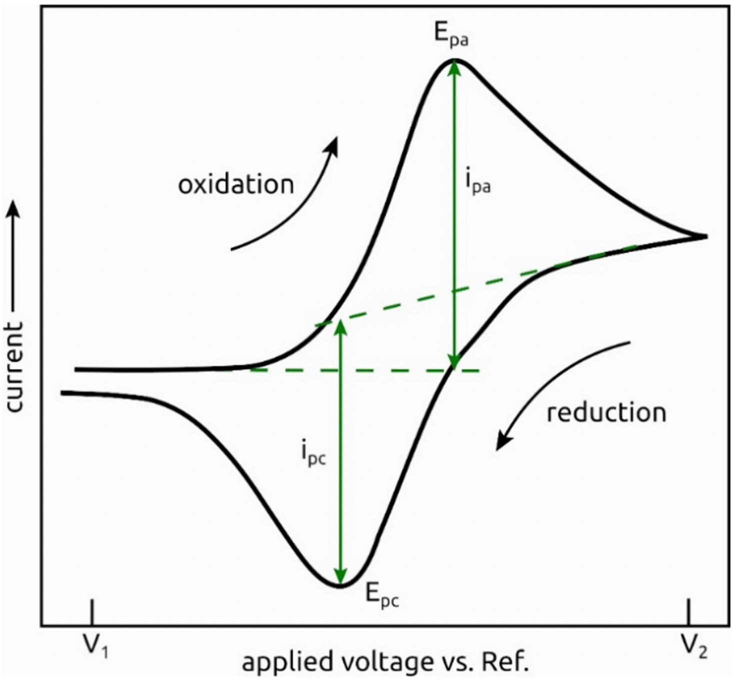


Figure 1.8*: Typical CV voltammogram. Image adapted from Kissinger*[56]

Through the incorporation of a redox tag on the recognition strand, nucleic acid sensing uses this technique to monitor the change in potential when hybridisation occurs between the target sequence and the nucleic acid conjugated to the surface of the working electrode.[57] Through the addition of a redox tag, such as methylene blue[58] or ferrocene,[59] at the terminal end of the nucleotide it is possible to monitor hybridisation (**Figure 1.9**). When no target is present the ssDNA can coil and fold to bring the redox tag close to the working electrode. Upon hybridisation the dsDNA possesses greater order, thus increasing the separation between the redox tag and the electrode surface, therefore resulting in a decrease in signal. This decrease in signal can be used to selectively identify the presence of a target sequence.[58,60]

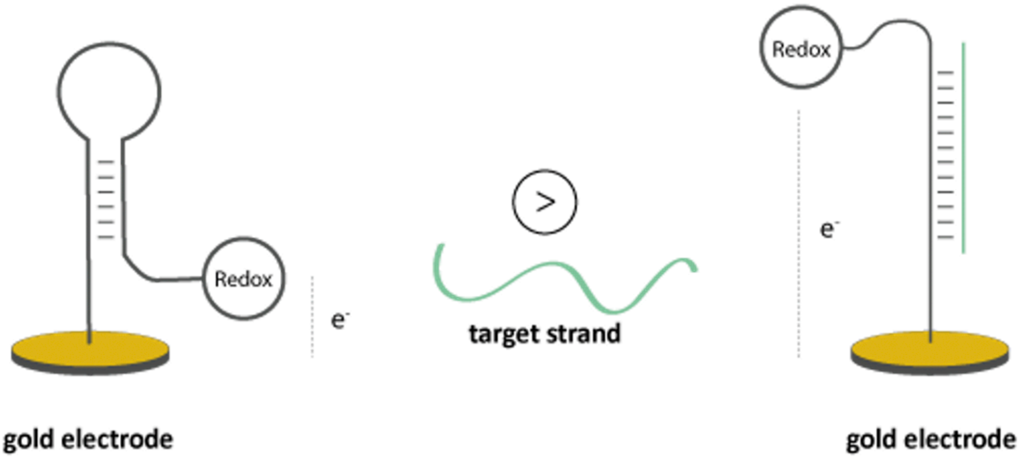


***Figure 1.9**: A nucleic acid probe with a redox tag. After hybridisation of the target strand there is a change in conformation increasing the distance between the redox tag with the electrode, thus leading to a decrease in signal. Image taken from Biomers.net.*[61]

Unfortunately, the samples for electrochemical sensing typically have to be clean samples with few ions in solution that can affect the result, as the greater the number of ion impurities the most likely a negative signal will arise.

1.4.3.2 Surface Plasmon Resonance

While electrochemical techniques are highly sensitive and capable of real time detection of DNA targets, they are not best suited to rapid diagnostics due to the requirement of a lengthy sample preparation. Therefore, a DNA sensor that is capable of real time detection of unpurfled DNA, thus improving samples preparation is favoured.[62] One such method that is capable of achieving this goal is surface plasmon resonance spectroscopy (SPR). SPR is a mass based, label free technique that is capable of monitoring real time hybridisation of DNA.[63] SPR eliminates the requirement of a label; which may affect the binding affinity of the target or introduce steric hinderance, thus increasing binding.[50,64,65] SPR monitors the interactions between free analytes in solution and fixed biorecognition elements bound to the surface of a sensor chip, which is traditionally made of gold.[64] A beam of light is coupled into a sensor chip through a prism, the angle in which the beam of light is reflected is monitored by a detector.[64–66] When a photon hits the metal surface of the detector chip, depending on the angle of incidence, some of the electrons at the metal surface are excited and capable of resonating, this is known as the resonance angle.. When the light source and metal surface are constant the angle in which resonance occurs is dependent on the refractive index of the sensing medium. A hybridisation event between biorecognition element and biomolecule, such as a target nucleic acid and the corresponding complementary sequence, at the surface results in a change in refractive index of the sensing medium, thus altering the resonance angle (**Figure 1.10**).

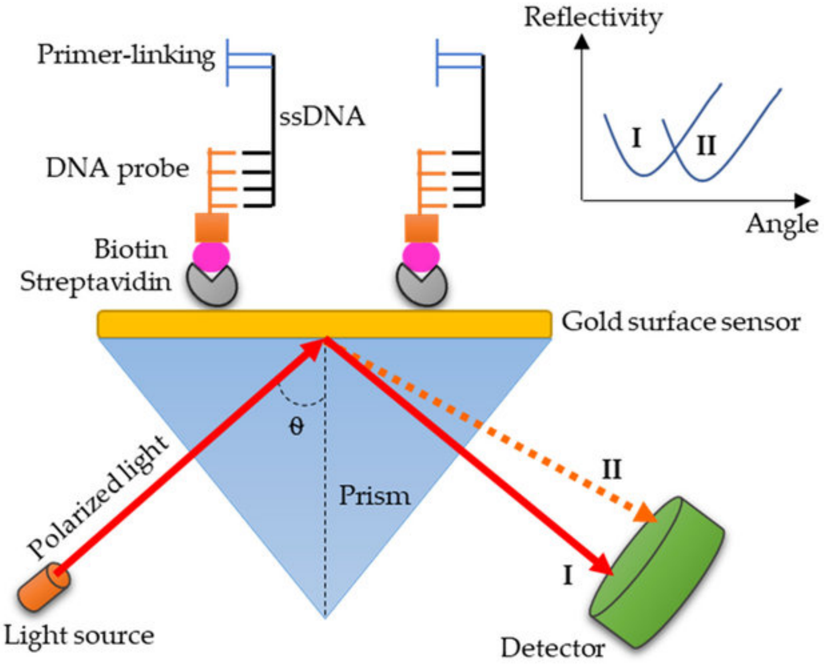


Figure 1.10: *SPR of DNA, (I) represents the resonance angle in the absence of ssDNA, (II) represents the resonance angle following a binding event at the surface of the sensor chip. Image taken from Alhaj-Qasem et al.* [67]

The change in resonance angles enables binding events to be monitored this is traditionally used for the detection of proteins;[68] however, recent developments have shown the incorporation of DNA sequences onto the chip.[69]

A drawback to DNA detection is the inability to differentiate between mutations, as a single base change can still result in the occurrence of hybridisation, thus generating a signal. A further problem which is being rapidly overcome is the high limit of detection, with conventional SPR instruments detecting DNA in the micromolar range which exceeds the concentration of genetical material of samples.[65,66,68]

1.4.3.3 Nanopore Sequencing

Nanopore sequencing has been widely deployed as a testing platform, specifically in West Africa during the Ebola outbreak by Professor Nick Loman from the University of Birmingham.[70] A nanopore is, as the name suggests, a pore that has a width of less than 1 nm and is usually a biological molecule; however, it can also be a hole that has been cut into the surface of a solid material. While solid state pores are less susceptible to changes in the environment (pH, temperature etc.) they are challenging to make and are not as controlled as biological molecules.[71] Cells are extremely good at creating nanopores which act as channels for transporting molecules into and out of a cell. These channels are found in the cell membrane surrounded by a lipid bilayer, which can be extracted and used in nanopore systems.[72,73] One such protein channel of particular interested for nanopore sequencing is α-hemolysin, a heptameric protein with a pore diameter of 1 nm, which is wide enough for a single strand of DNA to pass through (**Figure 1.11**).[74]

In order to create the nanopore sequencer,[75] an α-hemolysin and the surrounding lipid bilayer are extracted and then immersed in an electrolytic solution[76], where a membrane divides the solution into two separate chambers.[77] An electric field is applied to the system causing the ions in the electrolyte solution to move to their opposing electrode, and this electrophoresis effect can be exploited to also move charged molecules such as DNA.[78] The negatively charged DNA is attracted to the anode which is designed to be situated on the opposite site of the nanopore thus driving the DNA through the pore. Unfortunately, this process is far too rapid to determine any information from the sequence. Therefore, the DNA can be coupled to an enzyme which slows down transit. Using the coupled approach, a nanopore sequencer unwinds the dsDNA, passing a single strand through the pore base by base.[71,76] A nanopore sequence measures the ionic current generated by the flow of ions; as each of the four DNA bases are of a different size they impede different quantities of the ions in solution.

As a result, each base produces a different ionic current, and this signal can be converted into a real-time DNA sequence (**Figure 1.12**).[79–81]

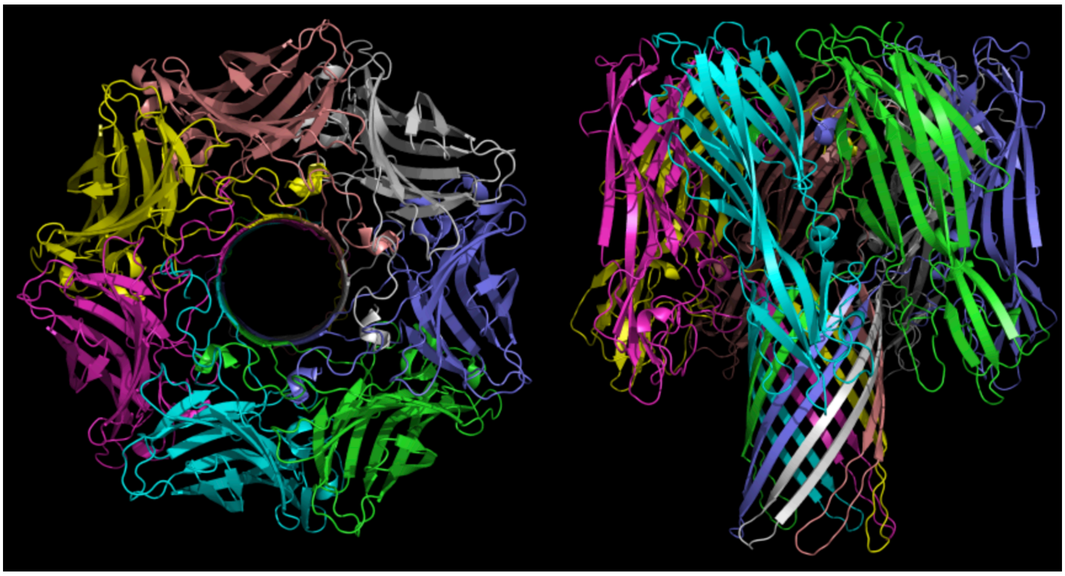


***Figure 1.11**: Crystal structure of α-hemolysin. Image adapted from PDB.*[82]

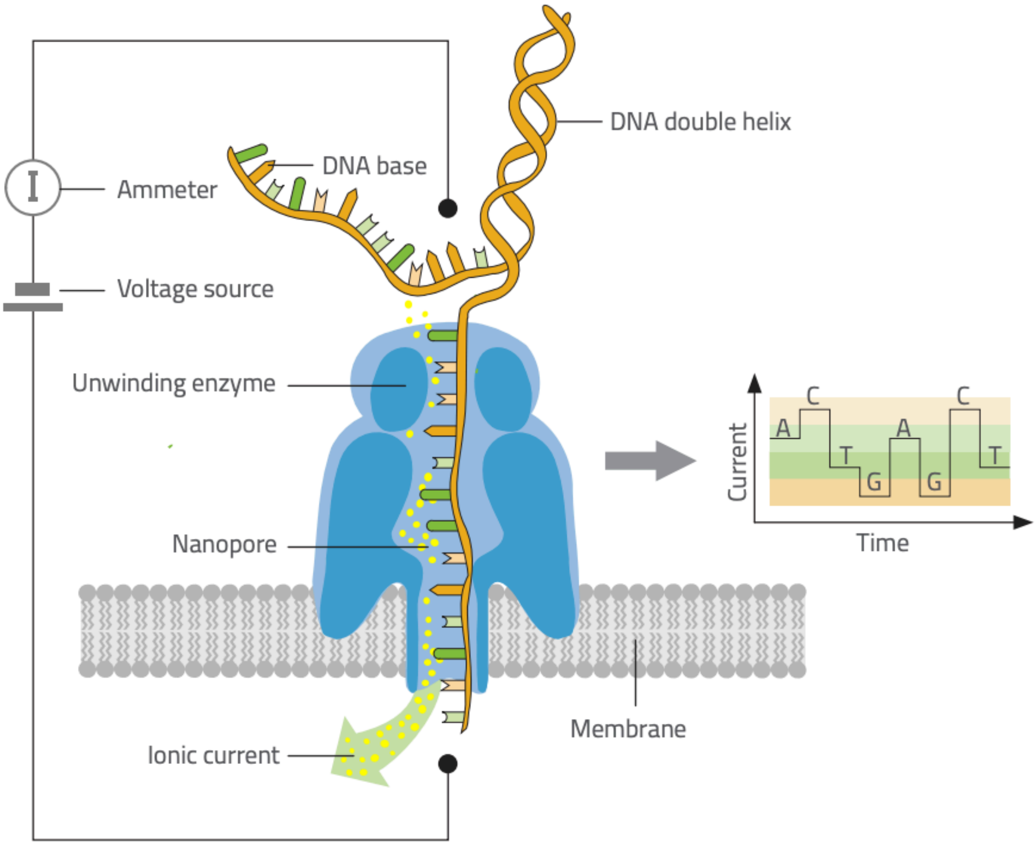


***Figure 1.12**: A representation of Nanopore. A dsDNA sequence is unwound by an enzyme, which feds a single strand into the nanopore. The different bases register a different current which can be used to determine the nucleic acid sequence. Image taken from Göpfrich.*[83]

Traditionally, DNA sequencing was not used as a method of detecting the presence of DNA. This was a result of the systems taking lengthy periods of time (usually days), being very expensive, or being far too bulky to be used outside of a specialised lab.[84] The advent of Oxford Nanopores minION sequencer enabled rapid, low cost sequencing in a handheld device, thus enabling rapid sequencing of viruses (Ebola and hepatitis C)[85] and detecting variants in cancer.[86] However, this system is not suited to rapidly differentiating between multiple sequences in a single sample.

1.4.3.4 Nucleic Acid Amplification

The final method that will be discusses is nucleic acid amplification; small quantities of specific DNA or RNA can be amplified to produce more copies of the starting nucleic acids, two of the most popular methods for amplifying nucleic acids are the polymerase chain reaction (PCR) and loop mediated isothermal amplification (LAMP).

1.4.3.4.1 Polymerase Chain Reaction

PCR is an amplification method that makes use of temperature variation to exponentially amplify double stranded sequences of DNA.[87] PCR is a process that requires a DNA polymerase enzyme and two oligonucleotide sequences called primers, which are designed to only hybridise at the 5′ end of both strands of the target DNA.[88] PCR itself is a three step process that has become a fundamental component of modern biochemistry; the three steps can be broken down into, denaturing, annealing and elongation.[89] A thermocycler can be used to vary the temperature at different stages of the reaction, thus creating ideal conditions for each step.

1) Stage one: denaturing, the temperature is raised to 95 °C to separate double stranded DNA (dsDNA) sequences into the two single stranded DNA (ssDNA) counterparts.[90,91]
2) Stage two: annealing, a decrease in temperature to between 50 – 60 °C enables the primers, which are significantly shorter than the full DNA sequence to bind to the single strands.
3) Stage three: elongation, an increase in temperature to 72 °C allows the DNA polymerase to extend the primers forming two identical sequences of dsDNA.

These three steps are cycled multiple times, exponentially creating more dsDNA of identical sequences (**Figure 1.13**).[90–92]

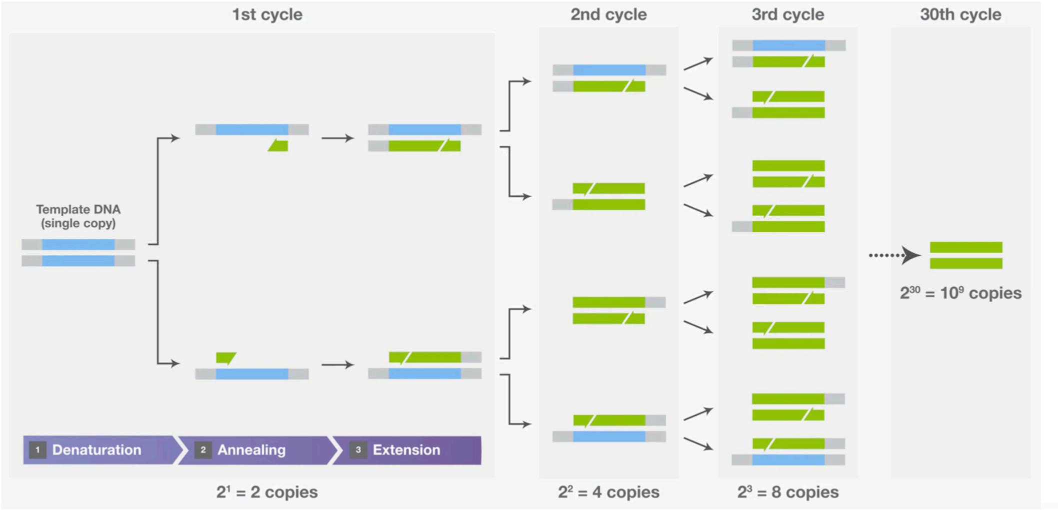


Figure 1.13: *The PCR process. During the first cycle the dsDNA is heat denaturing, separates the target into two single strands. A decrease in temperature allows primers to bind at the 3´ ends of the single strands. DNA polymerase extends the primers to create two new double stranded sequences of DNA. The cycle repeats, each cycle doubling the quantity of DNA. Image taken from ThermoFisher.*[93]

PCR is often referred to as the ‘gold standard’ as the development of PCR revolutionised the biochemistry world, enabling a tasks which previously took days to take hours. It is regularly used in most, if not all, biochemistry labs globally for cloning, sequencing genes and detecting pathogens, with minimal quantities of sample.[93,94] PCR made this possible due to its high specificity and sensitivity compared to most methods of detection, this coupled with the ease of primer design makes PCR an ideal method of diagnosing the cause of infection.[95]

1.4.3.4.1.1 Digital Droplet PCR

The application of PCR into diagnostic fields has enabled for the detection of many diseases where the concentration of genetic material is extremely low. An example being, the limit of detection for many human immunodeficiency virus tests is 50 copies/µL, this poses a problem when accurate quantification of <50 copies/µL is critical when guiding antiretroviral therapies.[96] Unfortunately, a drawback to PCR is the makeup of the sample is likely to contain cells, proteins and other sequences of DNA which can interfere and potentially even inhibit the detection of the target analyte.[97,98] While the favoured method to address this problem comes through sample clean up steps, they come with increased complexity while also reducing the concentration of the target analyte.[96,97] A novel method to avoid such problems is seen in the form of microfluidics, with the most sensitive method being digital droplet PCR (ddPCR).[99] ddPCR splits DNA molecules into approximately 20,000 droplets, with each droplet containing 1 copy of DNA. Through the introduction of PCR reagents amplification of a single copy can not only increase the amplification speed but also increase the sensitivity of the reaction.[100,101] In addition to the PCR reagents, magnetic beads for purification can be introduced, the amplified DNA can be extracted and subsequently tagged with a fluorescent antibody. Using flow cytometry it is the possible to individually count the number of fluorescent moieties thus reducing the likelihood of missing low copy number DNA in the background noise.[102–104] Unfortunately, despite the increases in speed and sensitivity ddPCR is an expensive technique, requiring more specialised equipment, not only for the amplification of single droplets of sample but also for the counting of amplicons.[99,103,104]

1.4.3.4.2 Loop Mediated Isothermal Amplification

Loop mediated isothermal amplification (LAMP) is an isothermal amplification method that relies on strand displacement activity.[105–107] LAMP is a highly sensitive and specific amplification method, that uses a strand displacement DNA polymerase, and four primers (two sets of inner primers and two sets of outer primers); the high sensitivity and specificity of LAMP come with the added cost of complexity and risk of laboratory contamination. The four LAMP primers are tailored to six regions of the target DNA, and as a result amplification can only occur when all six priming sites are present.[108] The target DNA is a double stranded sequence, containing six priming sites identified as F3c, F2c, F1c, B1c, B2c and B3c, the F sites are on one strand and the B sites on the complementary strand (**Figure 1.14**).[105,107]

Figure 1.14: *Double stranded target DNA with the priming sites labelled as F3c, F2c, F1c, B1c, B2c, and B3c with the complementary sequences also indicated.*

The four primers can be broken down in to two sets, the outer primers (OP) which are complementary to F3c and B3c and, as such, are designated as F3 and B3 respectively.[105–107] The inner primers (IP) are composed slightly differently. Looking at the forward inner primer (FIP), there is a sequence complementary to F1 at the 5´ end, designated F1c, linked to a sequence complementary to F2c, designated F2, at the 3´ end. These two sequences are separated by a poly T linker.[105,107] The backward inner primer (BIP) follows the same convention for the B sites (**Figure 1.15**).

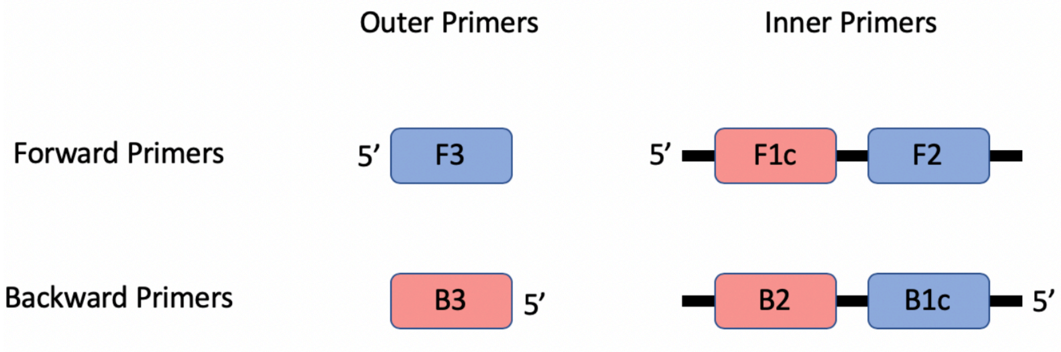


Figure 1.15: *The forward and backward, inner, and outer primers used within a LAMP reaction.*

Unlike PCR, LAMP is an isothermal amplification technique, in that the reaction all occurs at a constant temperature, this temperature is usually between 60 – 65 °C as this is the optimal temperature of the DNA polymerase used. The reaction beings when the F2 region of the FIP anneals to the F2c area of the target sequence, meanwhile the forward outer primer (FOP) hybridises to the F3c region. The DNA polymerases is able to then extend the FIP to create a new dsDNA sequence. Simultaneously the BIP and backwards outer primer (BOP) anneal to their complementary sequences. (**Figure 1.16**).[105,107,108]

Figure 1.16: *The first stage of a LAMP reaction, the inner forward primer binds to F2c. DNA polymerase amplifies the sequence. Meanwhile, the outer primer F3 is capable of binding to F3c.*

Following amplification of the inner primer, DNA polymerase synthesises a strand starting at the outer primer, this amplification process displaces the previous strand (**Figure 1.17**).

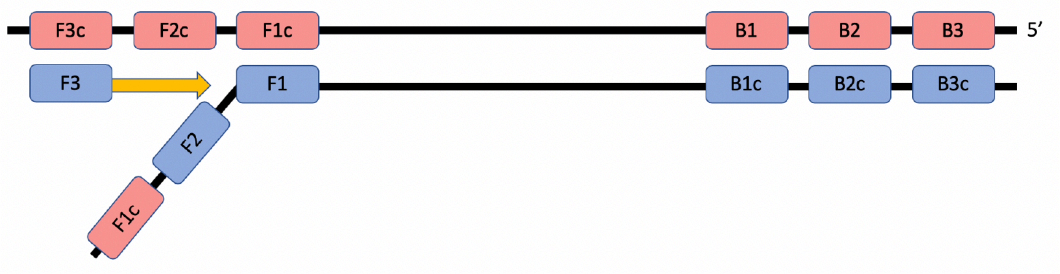


Figure 1.17: *Following amplification of the inner primer, DNA polymerase amplifies the outer primer, displacing the former strand in the process.*

The F1c region of the displaced strand forms a loop with the F1 area of the same sequence. The newly synthesised strand additionally possesses priming sties for the backwards primers, the same process repeated from these sites, leading to the formation of a 'dumb bell' like structure (**Figure 1.18**).[105,107] Following the formation of the 'dumb-bell' structure further primers are capable of binding; this priming leads to the formation of an extended structure through further amplification, displacement and priming steps (**Figure 1.18**).[105–107] The final product is a mix of stem-loops of varying length and cauliflower-like structures with multiple loops formed through annealing of inverted loops of target in the same sequence. The formation of these structures increases the turbidity of the solution resulting in a visible indicator that the reaction has occurred. However, in order to quantify the amplification various fluorophores can be introduced into the reaction.[105,107,109] While LAMP is becoming increasingly popular, it is not the only isothermal method of amplifying DNA, further methods will be discussed in (**Chapter 3.2**).

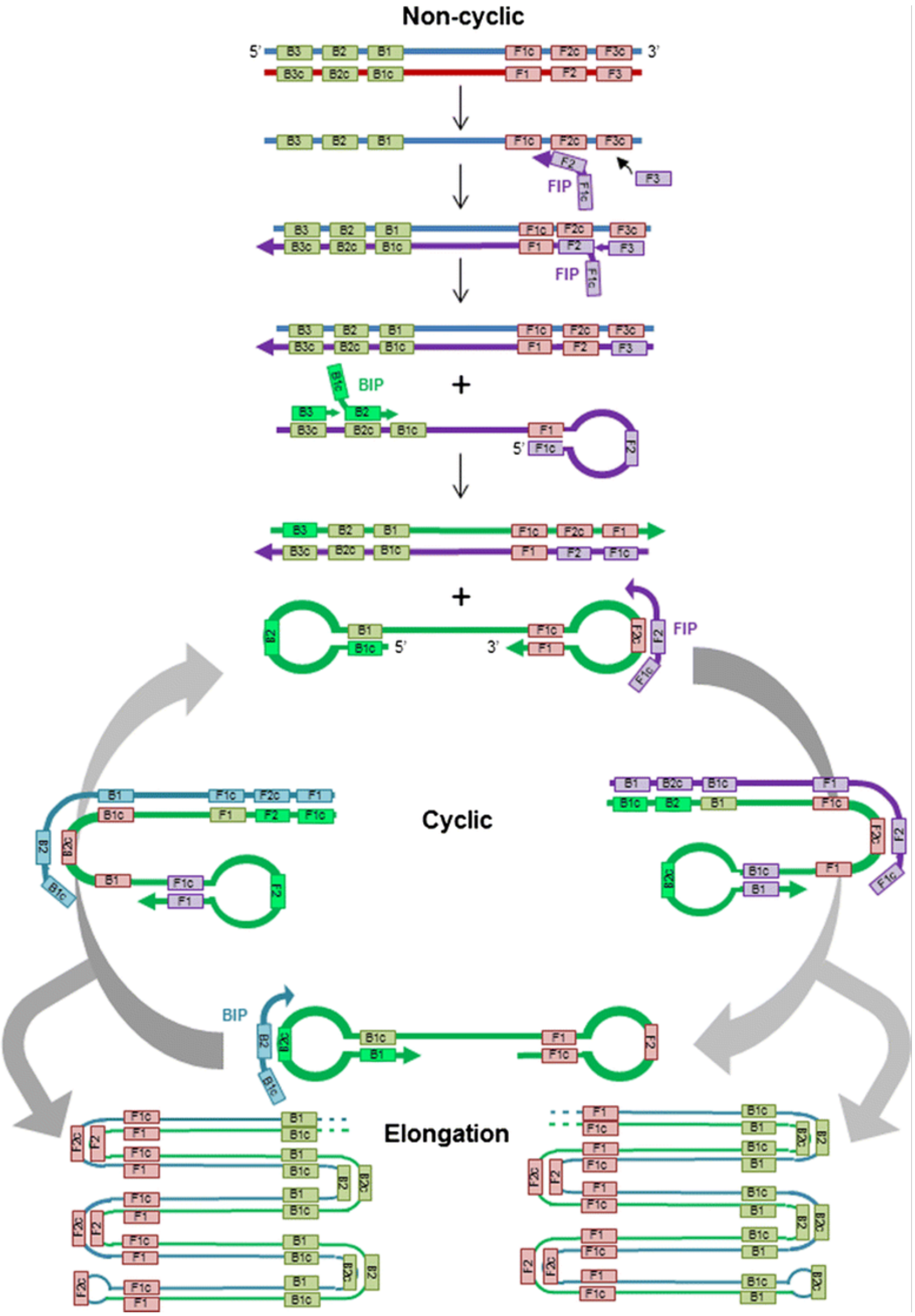


Figure 1.18: *A schematic depicting the LAMP process. Image taken from Bruce et al.*[110]

1.4.3.5 Detection of Nucleic Acid Amplification

The amplification of specific sequence of genetic material can be monitored using various methods. Some of the most adopted methods for detecting amplification will be discussed below.

1.4.3.5.1 Intercalating dyes

When a molecule absorbs light of an appropriate frequency, electrons are excited to a higher energy state. The excited state only exists for a matter of nanoseconds, during which time energy is dissipated through internal conversions. The electron returns to the ground state, emitting a photon of energy equal to the new energy gap in the form of fluorescence.[111,112] The decrease in energy results in a different wavelength being emitted, with the difference in excitation and emission wavelength being called the Stokes' shift.[111–113] Monitoring the transmittance of light at the emission wavelength enables a detection technique for the identification of fluorescent moieties. The use of fluorescent dyes is currently the most common method for identification of amplification. Real-time monitoring of DNA amplification, usually through PCR or LAMP can be performed using a fluorescent dye, such as SYBR Green I (**Figure 1.19**), which is capable of intercalating between the bases of DNA, offering a simple yet effective method for monitoring the production of target DNA.

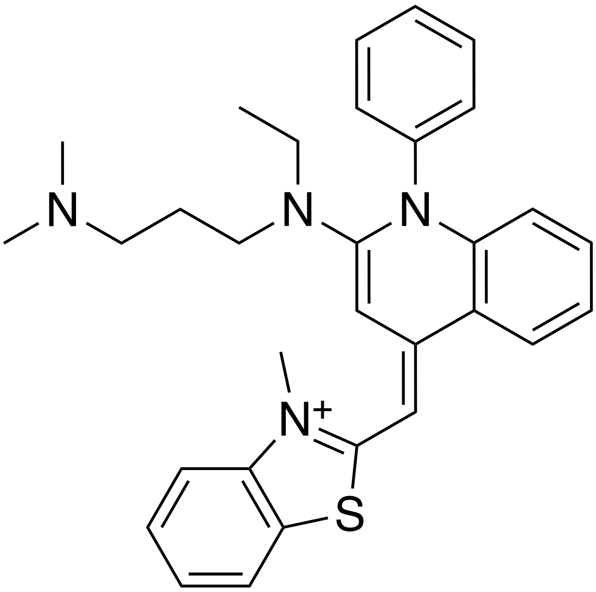


***Figure 1.19**: structure of the fluorescent intercalating dye, SYBR Green I.*

In nucleic acid amplification tests (NAATs) amplification can only occur when the target DNA is present; amplification produces dsDNA, into which the fluorescent moiety is able to intercalate, resulting in the fluorescent output of the dye increasing ~1000 fold.[114–117] However, due to the non-specificity of these dyes all amplification will produce an increase in fluorescence, thus rendering the ability to differentiate between different targets in the same sample moot. Consequentially sequence specific oligonucleotide probes are favoured with fluorescent dyes incorporated into the sequence. Binding of these probes results in an increase in fluorescence confirming the presence of target DNA.[118]

1.4.3.5.2 Molecular Beacons

Molecular beacons (MB) for nucleic acid detection are composed of oligonucleotides with a fluorophore on one end and a quencher at the other.[119] The oligonucleotide sequence itself contains a region that is fully complementary to the target sequence of choice, flanked by two regions that are complementary to each other; this allows the MB form a hairpin structure when the target is not present (**Figure 1.20**). The formation of the hairpin structure brings the fluorophore and quencher into close proximity of each other, where the quencher is subsequently able to quench the fluorescent emission form the fluorophore.[119,120] Upon introduction of the target, the MB changes conformation to allow the centre sequence, which is fully complementary to the target, to hybridise. This is driven by thermodynamic stability, as the longer centre sequence is more stable than the self-complementary region. The change in conformation leads to the fluorophore and quencher now being separated, therefore the quencher is unable to hold back the emission, and as a result the MB 'lights up'.[119,120] The fluorescent emission is proportional to the quantity of target in solution; the more target, the greater the fluorescent output, and this can be detected by a fluorimeter. Unlike detection *via* intercalating dye this is a sequence specific technique, however samples with small quantities of genetic material are more difficult to detect.

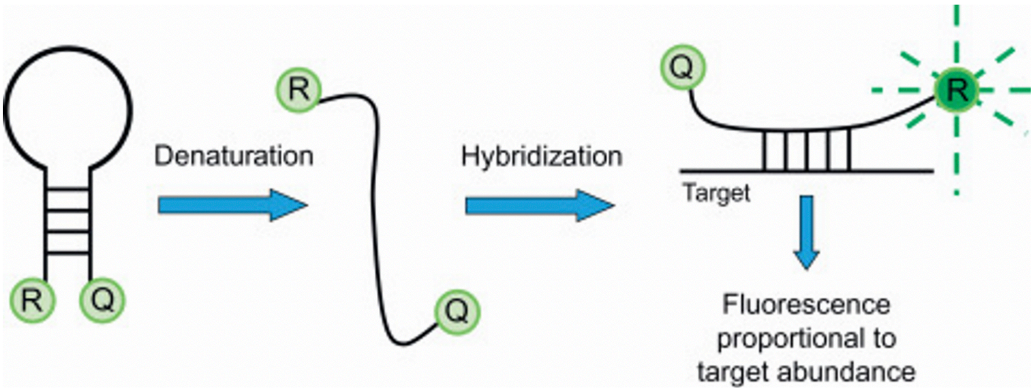


***Figure 1.20**: A representation of a molecular beacon, when the target is not present the self-complementary sequences hybridise bringing the fluorophore (R) and quencher (Q) into close proximity. Upon introduction of the target sequence the change in conformation leaves R unquenched and allowed to fluoresce. Image taken from Farrell.*[119]

A MB that has been designed specifically for using in a PCR reaction is the hydrolysis probe TaqMan. As with a standard MB, a TaqMan probe is composed of an oligonucleotide with a fluorophore and quencher at opposite ends, however it is lacking the self-complementary sequence.[121] The oligonucleotides are designed to have the fluorophore and quencher in close proximity for the emission to be quenched. For TaqMan to work a DNA polymerase from the bacterial species *Thermus aquaticus*, hence Taq, must be employed.[122] During the PCR reaction dsDNA is denatured to enable primer binding, in order for amplification to occur; the TaqMan probe is also able to hybridise at this time with the target sequence to which it is fully complementary. As the DNA polymerase extends the forward primer the exonuclease activity of the DNA polymerase degrades the TaqMan probe, liberating the fluorophore (**Figure 1.21**). The degradation results in the fluorophore and the quencher no longer being in close enough proximity so an increase in fluorescence can be observed; each successive cycle leads to more fluorophores being liberated and more fluorescence being recorded.[121,123–125]

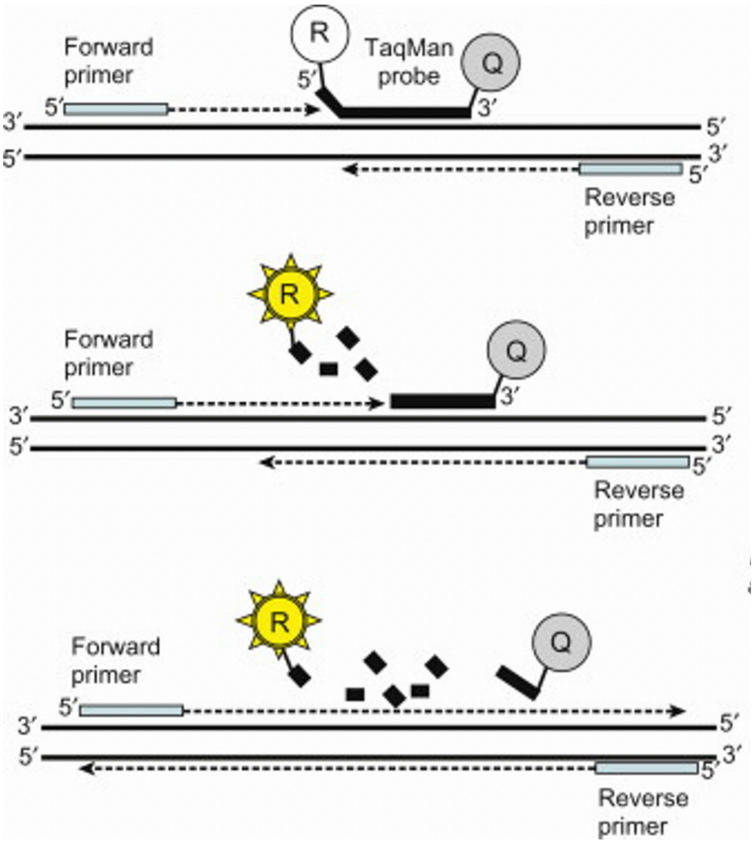


***Figure 1.21**: A representation of a TaqMan Assay. The TaqMan probe anneals with the PCR target, as does the PCR primers. Extension of the primer leads to the TaqMan probe being degraded and the fluorophore (R) being released. The cleavage of R from proximity of the quencher (Q) enables fluorescence emission to occur. Image adapted from Butler.*[126]

Fluorescent readouts are currently the most widely employed method of detecting the presence of nucleic acids, however they predominately rely on PCR amplification. While this is the ‘gold standard’ of DNA detection, it is a complex system that requires trained personnel to operate and interpret results, regardless of the simplicity of the detection method. Furthermore, the cost and size of a thermocycler are a hinderance for rapid diagnosis. While fluorescent probes are an ideal detection method, they still require further optimisation to be incorporated into isothermal amplification techniques, the future of nucleic acid detection

1.4.4 Linear Dichroism Sensing

In addition to the above methods there is an alternative method that will be a major method of detection in this thesis. Linear Dichroism (LD) is an optical absorbance technique that measures the alignment of a sample. Previous work within the Tucker and Dafforn groups, has led to the development of a LD based detection system for the detection of nucleic acids using M13 bacteriophage.[127–129] Bacteriophage are a group of viruses which are able to infect bacteria, one of these species is M13 bacteriophage. M13 bacteriophage is a filamentous bacteriophage, therefore possesses a long, thin, rod-like structure composed of DNA encapsulated in a protein coat.[130,131] (**Figure 1.22**).

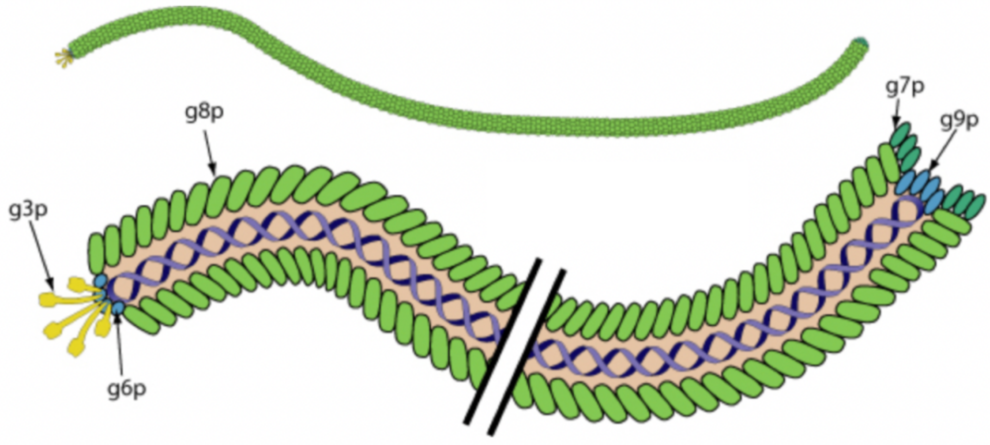


***Figure 1.22**: An example of a filamentous bacteriophage.*[132]

The protein coat is constructed of 5 protein species of pIII, pVI, pVII, pVIII and pIX;[133] which assemble to make the M13 bacteriophage approximately 900 nm long and 7 nm in diameter as a result this bacteriophage possesses a very large aspect ratio.[134] The high aspect ratio, makes M13 bacteriophage amenable to study *via* LD. When subjected to shear flow the M13 align along their long axis in the direction of flow, thus giving rise to a large distinct LD signal (**Figure 1.25**).[127–129,135–137] The major coat protein of M13 possess an external free lysine residue, thus a terminal amine; by using the heterobifunctional crosslinker, succinimidyl 4-(N-maleimidomethyl) cyclohexane-1-carboxylate (SMCC) it was possible to conjugate DNA and M13 (**Figure 1.23**).

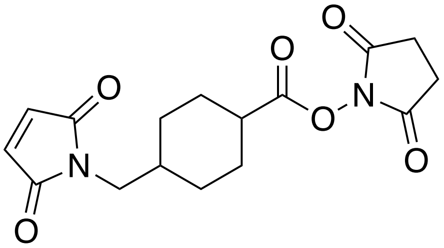


Figure 1.23: Structure of succinimidyl 4-(N-maleimidomethyl) cyclohexane-1-carboxylate.

The free amine from the lysine residue reacts with the *N*-hydroxysuccinimide (*NHS)* ester forming an amide bond. DNA probes were synthesised with a thiol tag, which was able to react with the maleimide group, cross linking the DNA to the M13. In order to create a DNA probe system a target sequence was selected, and two complementary sequences were designed to be complementary to half the sequence each, the two probes combined being complementary to the full target sequence. The inclusion of a thiol tag at the 5´ end of one sequence, and the 3´ end of the other enabled a sandwich assay to be created. In the presence of the target DNA strand, hybridisation occurred, causing the M13-probes to bind together (**Figure 1.24**).[129,138,139]

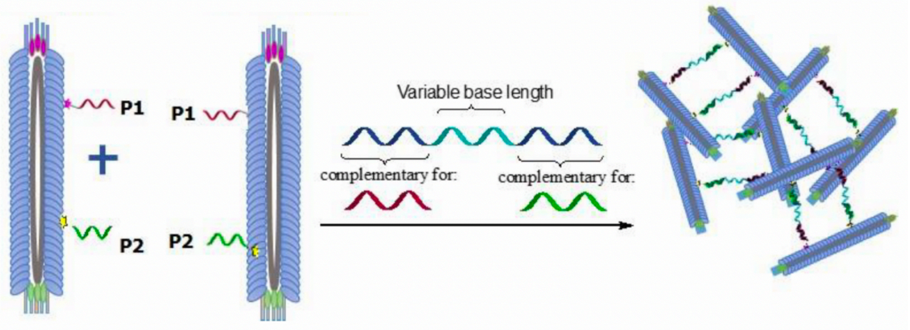


***Figure 1.24**: The 2 probe, 1 phage sandwich assay. The left hand shows the M13-probes without target present; the right hand shows the M13-probes when the correct target is present. Image taken from Little.*[138]

The linking of multiple M13 molecules results in aggregation; when the probes are unbound they are able to align in the same direction, following aggregation the M13 molecules can no longer align in flow, and as a result there is a decrease in LD signal (**Figure 1.25**).[129,138,139] The phage system was developed into an assay to detect the distinctive base sequences of the genes that code for the coat proteins of PVA and PVY potato viruses. This assay was capable of detecting 50 pM of genetic material.[129,139]

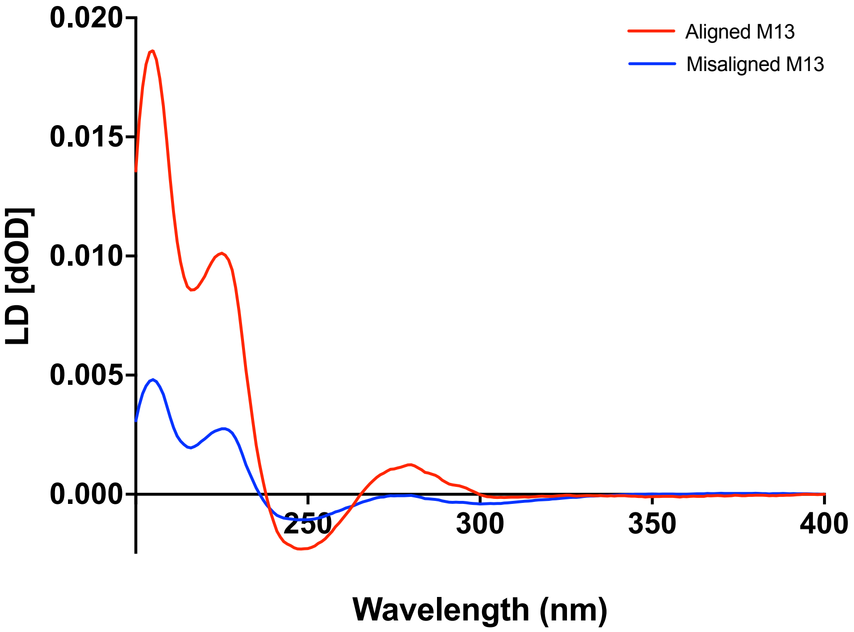


***Figure 1.25**: LD spectra displaying the signal generated from aligned M13 when target DNA is not present (Red) and misaligned M13 when target DNA is present (Blue).*

1.5 Thesis Outline

This thesis details the work of three projects carried out during this PhD based upon the development of a novel amplification method for RNA and DNA and the exploration of a new reagent for LD assays based around foldamers.

Chapter 2 outlines the theory behind the techniques used during this research. This chapter features a detailed discussion of dichroism spectroscopy.

Chapter 3 discusses the optimisation of a rapid exponential amplification reaction, for the production of DNA that can be detected using an established linear dichroism-based bioassay. This chapter will explore nucleic acid amplification techniques and investigate linear dichroism based M13 bioassays developed by the group.

Chapter 4 discusses the development of a rapid nucleic acid-based assay for the detection of SARS-CoV-2 RNA. This chapter will discuss current testing strategies and the trials and tribulations attributed to developing a novel method of detection during a global pandemic.

Chapter 5 explores the linear dichroism signals produced from polymeric molecules which change conformation based upon their environment and assesses their ability to form a novel scaffold for LD based DNA detection. This chapter will detail the nature of foldamers, their origins and uses.

Chapter 6 details the methods used throughout this project.

1.6 References

1. Fleischmann, C. *et al.* Hospital Incidence and Mortality Rates of Sepsis. *Dtsch Arztebl Int* **113**, 159–166 (2016).
2. Strimbu, K. & Tavel, J. A. What are biomarkers?: *Current Opinion in HIV and AIDS* **5**, 463–466 (2010).
3. Mak, W. C., Beni, V. & Turner, A. P. F. Lateral-flow technology: From visual to instrumental. *TrAC Trends in Analytical Chemistry* **79**, 297–305 (2016).
4. Koczula, K. M. & Gallotta, A. Lateral flow assays. *Essays In Biochemistry* **60**, 111–120 (2016).
5. You, M. *et al.* Household Fluorescent Lateral Flow Strip Platform for Sensitive and Quantitative Prognosis of Heart Failure Using Dual-Color Upconversion Nanoparticles. *ACS Nano* **11**, 6261–6270 (2017).
6. Pilavaki, E. & Demosthenous, A. Optimized Lateral Flow Immunoassay Reader for the Detection of Infectious Diseases in Developing Countries. *Sensors* **17**, 2673 (2017).
7. Kim, C. *et al.* Battery operated preconcentration-assisted lateral flow assay. *Lab Chip* **17**, 2451–2458 (2017).
8. Naidoo, N. *et al.* Modified RS-LAMP assay and use of lateral flow devices for rapid detection of Leifsonia xyli subsp. xyli. *Lett Appl Microbiol* **65**, 496–503 (2017).
9. Xiao, W. *et al.* A simple and compact smartphone-based device for the quantitative readout of colloidal gold lateral flow immunoassay strips. *Sensors and Actuators B: Chemical* **266**, 63–70 (2018).
10. Nelis *et al.* The Efficiency of Color Space Channels to Quantify Color and Color Intensity Change in Liquids, pH Strips, and Lateral Flow Assays with Smartphones. *Sensors* **19**, 5104 (2019).
11. Schwenke, K. U. *et al.* Analysis of free chlorine in aqueous solution at very low concentration with lateral flow tests. *Sci Rep* **9**, 17212 (2019).

12. Lee, S., Mehta, S. & Erickson, D. Two-Color Lateral Flow Assay for Multiplex Detection of Causative Agents Behind Acute Febrile Illnesses. *Anal. Chem.* **88**, 8359–8363 (2016).
13. Yen, C.-W. *et al.* Multicolored silver nanoparticles for multiplexed disease diagnostics: distinguishing dengue, yellow fever, and Ebola viruses. *Lab Chip* **15**, 1638–1641 (2015).
14. Byzova, N. *et al.* Lateral Flow Immunoassay for Rapid Detection of Grapevine Leafroll-Associated Virus. *Biosensors* **8**, 111 (2018).
15. Borges, M. A. S. B., Araújo Filho, J. A. de, Soares, R. de B. A., Vidal, J. E. & Turchi, M. D. False-negative result of serum cryptococcal antigen lateral flow assay in an HIV-infected patient with culture-proven cryptococcaemia. *Medical Mycology Case Reports* **26**, 64–66 (2019).
16. Foysal, K. H., Seo, S. E., Kim, M. J., Kwon, O. S. & Chong, J. W. Analyte Quantity Detection from Lateral Flow Assay Using a Smartphone. *Sensors* **19**, 4812 (2019).
17. Romeo, A., Leung, T. S. & Sánchez, S. Smart biosensors for multiplexed and fully integrated point-of-care diagnostics. *Lab Chip* **16**, 1957–1961 (2016).
18. Rohrman, B. A., Leautaud, V., Molyneux, E. & Richards-Kortum, R. R. A Lateral Flow Assay for Quantitative Detection of Amplified HIV-1 RNA. *PLoS ONE* **7**, e45611 (2012).
19. Kamphee, H. *et al.* Rapid Molecular Detection of Multidrug-Resistant Tuberculosis by PCR-Nucleic Acid Lateral Flow Immunoassay. *PLoS ONE* **10**, e0137791 (2015).
20. Gitonga, L. K. *et al.* Point of care testing evaluation of lateral flow immunoassay for diagnosis of cryptococcus meningitis in HIV-positive patients at an urban hospital in Nairobi, Kenya, 2017. *BMC Res Notes* **12**, 797 (2019).
21. Chen, Z. *et al.* Rapid and Sensitive Detection of anti-SARS-CoV-2 IgG, Using Lanthanide-Doped Nanoparticles-Based Lateral Flow Immunoassay. *Anal. Chem.* **92**, 7226–7231 (2020).
22. McPartlin, D. A. & O'Kennedy, R. J. Point-of-care diagnostics, a major opportunity for change in traditional diagnostic approaches: potential and limitations. *Expert Review of Molecular Diagnostics* **14**, 979–998 (2014).

23. Fu, X. *et al.* Highly sensitive detection of prostate cancer specific PCA3 mimic DNA using SERS-based competitive lateral flow assay. *Nanoscale* **11**, 15530–15536 (2019).
24. Lu, F., Wang, K. H. & Lin, Y. Rapid, quantitative and sensitive immunochromatographic assay based on stripping voltametric detection of a metal ion label. *Analyst* **130**, 1513 (2005).
25. Cross, R. W. *et al.* Analytical Validation of the ReEBOV Antigen Rapid Test for Point-of-Care Diagnosis of Ebola Virus Infection. *J Infect Dis.* **214**, S210–S217 (2016).
26. Nouvellet, P. *et al.* The role of rapid diagnostics in managing Ebola epidemics. *Nature* **528**, S109–S116 (2015).
27. Travers, A. & Muskhelishvili, G. DNA structure and function. *FEBS J* **282**, 2279–2295 (2015).
28. Kabanova, S. *et al.* Gene expression analysis of human red blood cells. *Int. J. Med. Sci.* 156–159 (2009) doi:10.7150/ijms.6.156.
29. What is DNA? - Genetics Home Reference - NIH. https://ghr.nlm.nih.gov/primer/basics/dna.
30. Sinden, R. R. *DNA structure and function*. (Academic Press, 1994).
31. *Molecular biology of the cell*. (Garland Science, 2002).
32. *Molecular cell biology*. (W.H. Freeman, 2000).
33. Harding, S. E., Channell, G. & Phillips-Jones, M. K. The discovery of hydrogen bonds in DNA and a re-evaluation of the 1948 Creeth two-chain model for its structure. *Biochemical Society Transactions* **46**, 1171–1182 (2018).
34. Burge, S., Parkinson, G. N., Hazel, P., Todd, A. K. & Neidle, S. Quadruplex DNA: sequence, topology and structure. *Nucleic Acids Research* **34**, 5402–5415 (2006).
35. Clancy, S. & Brown, W. Translation: DNA to mRNA to Protein. *Nature Education* **1**, 101 (2008).
36. *Biochemistry*. (Pearson, 2013).
37. Holley, R. W. *et al.* Structure of a Ribonucleic Acid. *Science* **147**, 1462–1465 (1965).
38. *Burger’s medicinal chemistry, drug discovery and development*. (Wiley, 2021).
39. Elliott, D. & Ladomery, M. *Molecular biology of RNA*. (Oxford University Press, 2016).

40. Mishanina, T. V., Koehn, E. M. & Kohen, A. Mechanisms and inhibition of uracil methylating enzymes. *Bioorganic Chemistry* **43**, 37–43 (2012).
41. Vértessy, B. G. & Tóth, J. Keeping Uracil Out of DNA: Physiological Role, Structure and Catalytic Mechanism of dUTPases. *Acc. Chem. Res.* **42**, 97–106 (2009).
42. Tinoco, I. & Bustamante, C. How RNA folds. *Journal of Molecular Biology* **293**, 271–281 (1999).
43. Li, J. & Liu, C. Coding or Noncoding, the Converging Concepts of RNAs. *Front. Genet.* **10**, 496 (2019).
44. Lin, S.-L., Miller, J. D. & Ying, S.-Y. Intronic MicroRNA (miRNA). *Journal of Biomedicine and Biotechnology* **2006**, 1–13 (2006).
45. Ahmad, K. & Henikoff, S. Epigenetic Consequences of Nucleosome Dynamics. *Cell* **111**, 281–284 (2002).
46. Woodhams, M. D., Stadler, P. F., Penny, D. & Collins, L. J. RNase MRP and the RNA processing cascade in the eukaryotic ancestor. *BMC Evol Biol* **7**, S13 (2007).
47. Litwack, G. Protein Biosynthesis. in *Human Biochemistry* 319–336 (Elsevier, 2018). doi:10.1016/B978-0-12-383864-3.00011-9.
48. Petrov, A. S. *et al.* Secondary Structures of rRNAs from All Three Domains of Life. *PLoS ONE* **9**, e88222 (2014).
49. Elgrishi, N. *et al.* A Practical Beginner’s Guide to Cyclic Voltammetry. *J. Chem. Educ.* **95**, 197–206 (2018).
50. Nahir, T. M., Clark, R. A. & Bowden, E. F. Linear-Sweep Voltammetry of Irreversible Electron Transfer in Surface-Confined Species Using the Marcus Theory. *Anal. Chem.* **66**, 2595–2598 (1994).
51. Ramaley, Louis. & Krause, M. S. Theory of square wave voltammetry. *Anal. Chem.* **41**, 1362–1365 (1969).

52. Nazari-Vanani, R., Sattarahmady, N., Yadegari, H. & Heli, H. A novel and ultrasensitive electrochemical DNA biosensor based on an ice crystals-like gold nanostructure for the detection of Enterococcus faecalis gene sequence. *Colloids and Surfaces B: Biointerfaces* **166**, 245–253 (2018).
53. Irkham *et al.* Co-reactant-on-Demand ECL: Electrogenerated Chemiluminescence by the in Situ Production of $S_2O_8{}^{2-}$ at Boron-Doped Diamond Electrodes. *J. Am. Chem. Soc.* **138**, 15636–15641 (2016).
54. Benvidi, A., Tezerjani, M. D., Jahanbani, S., Mazloum Ardakani, M. & Moshtaghioun, S. M. Comparison of impedimetric detection of DNA hybridization on the various biosensors based on modified glassy carbon electrodes with PANHS and nanomaterials of RGO and MWCNTs. *Talanta* **147**, 621–627 (2016).
55. Pacchioni, G. A not-so-strong bond. *Nat Rev Mater* **4**, 226–226 (2019).
56. Kissinger, P. T. & Heineman, W. R. Cyclic voltammetry. *J. Chem. Educ.* **60**, 702 (1983).
57. Gaudin, V. Advances in biosensor development for the screening of antibiotic residues in food products of animal origin – A comprehensive review. *Biosensors and Bioelectronics* **90**, 363–377 (2017).
58. Otero, F., Shortall, K., Salaj-Kosla, U., Tofail, S. A. M. & Magner, E. Electrochemical biosensor for the detection of a sequence of the TP53 gene using a methylene blue labelled DNA probe. *Electrochimica Acta* **388**, 138642 (2021).
59. Duprey, J.-L. H. A. & Tucker, J. H. R. Metal–Carbon Bonds in Biopolymer Conjugates: Bioorganometallic Nucleic Acid Chemistry. *Chem. Lett.* **43**, 157–163 (2014).
60. Lubin, A. A. & Plaxco, K. W. Folding-Based Electrochemical Biosensors: The Case for Responsive Nucleic Acid Architectures. *Acc. Chem. Res.* **43**, 496–505 (2010).
61. biomers.net. Oligonucleotides for Electrochemical Detection. https://www.biomers.net/en/products/dna/electrochemical_detection.html.

62. Sakao, Y., Nakamura, F., Ueno, N. & Hara, M. Hybridization of oligonucleotide by using DNA self-assembled monolayer. *Colloids and Surfaces B: Biointerfaces* **40**, 149–152 (2005).
63. Kawata, S. Plasmonics for Nanoimaging and Nanospectroscopy. *Appl Spectrosc* **67**, 117–125 (2013).
64. Homola, J., Yee, S. S. & Gauglitz, G. Surface plasmon resonance sensors: review. *Sensors and Actuators B: Chemical* **54**, 3–15 (1999).
65. Gambari, R. & Feriotto, G. Surface Plasmon Resonance for Detection of Genetically Modified Organisms in the Food Supply. *Journal of AOAC INTERNATIONAL* **89**, 893–897 (2006).
66. Gong, P., Lee, C.-Y., Gamble, L. J., Castner, D. G. & Grainger, D. W. Hybridization Behavior of Mixed DNA/Alkylthiol Monolayers on Gold: Characterization by Surface Plasmon Resonance and ^{32} P Radiometric Assay. *Anal. Chem.* **78**, 3326–3334 (2006).
67. Alhaj-Qasem, D. M. *et al.* Laboratory Diagnosis of Paratyphoid Fever: Opportunity of Surface Plasmon Resonance. *Diagnostics* **10**, 438 (2020).
68. Feriotto, G., Breveglieri, G., Finotti, A., Gardenghi, S. & Gambari, R. Real-time multiplex analysis of four beta-thalassemia mutations employing surface plasmon resonance and biosensor technology. *Lab Invest* **84**, 796–803 (2004).
69. Vasen, H. F. A. *et al.* Early detection of breast and ovarian cancer in families with BRCA mutations. *European Journal of Cancer* **41**, 549–554 (2005).
70. Stopping outbreaks becoming epidemics. *University of Birmingham* https://www.birmingham.ac.uk/research/quest/emerging-frontiers/a-lab-in-a-suitcase.aspx.
71. Types of nanopores. *Oxford Nanopore Technologies* http://nanoporetech.com/how-it-works/types-of-nanopores.
72. Grime, R. L. *et al.* Differences in SMA-like polymer architecture dictate the conformational changes exhibited by the membrane protein rhodopsin encapsulated in lipid nano-particles. *Nanoscale* **13**, 13519–13528 (2021).

73. Grime, R. L. *et al.* Single molecule binding of a ligand to a G-protein-coupled receptor in real time using fluorescence correlation spectroscopy, rendered possible by nano-encapsulation in styrene maleic acid lipid particles. *Nanoscale* **12**, 11518–11525 (2020).
74. Kasianowicz, J. J., Brandin, E., Branton, D. & Deamer, D. W. Characterization of individual polynucleotide molecules using a membrane channel. *Proceedings of the National Academy of Sciences* **93**, 13770–13773 (1996).
75. MinION. *Oxford Nanopore Technologies* http://nanoporetech.com/products/minion.
76. Varongchayakul, N., Song, J., Meller, A. & Grinstaff, M. W. Single-molecule protein sensing in a nanopore: a tutorial. *Chem. Soc. Rev.* **47**, 8512–8524 (2018).
77. Talaga, D. S. & Li, J. Single-Molecule Protein Unfolding in Solid State Nanopores. *J. Am. Chem. Soc.* **131**, 9287–9297 (2009).
78. Chen, P. *et al.* Probing Single DNA Molecule Transport Using Fabricated Nanopores. *Nano Lett.* **4**, 2293–2298 (2004).
79. Lu, H., Giordano, F. & Ning, Z. Oxford Nanopore MinION Sequencing and Genome Assembly. *Genomics, Proteomics & Bioinformatics* **14**, 265–279 (2016).
80. Tyler, A. D. *et al.* Evaluation of Oxford Nanopore's MinION Sequencing Device for Microbial Whole Genome Sequencing Applications. *Sci Rep* **8**, 10931 (2018).
81. Loman, N. J., Quick, J. & Simpson, J. T. A complete bacterial genome assembled de novo using only nanopore sequencing data. *Nat Methods* **12**, 733–735 (2015).
82. Song, L. *et al.* Structure of Staphylococcal α-Hemolysin, a Heptameric Transmembrane Pore. *Science* **274**, 1859–1865 (1996).
83. Decoding DNA with a pocket-sized sequencer – Science in School. https://www.scienceinschool.org/article/2018/decoding-dna-pocket-sized-sequencer/.
84. Liu, L. *et al.* Comparison of Next-Generation Sequencing Systems. *Journal of Biomedicine and Biotechnology* **2012**, 1–11 (2012).

85. Greninger, A. L. *et al.* Rapid metagenomic identification of viral pathogens in clinical samples by real-time nanopore sequencing analysis. *Genome Med* **7**, 99 (2015).
86. Norris, A. L., Workman, R. E., Fan, Y., Eshleman, J. R. & Timp, W. Nanopore sequencing detects structural variants in cancer. *Cancer Biology & Therapy* **17**, 246–253 (2016).
87. National Laboratory of Enteric Pathogens, Bureau of Microbiology, Laboratory Centre for Disease Control. The polymerase chain reaction: An overview and development of diagnostic PCR protocols at the. LCDC. *The Canadian Journal of Infectious Disease* **2**, 88–91 (1991).
88. Garibyan, L. & Avashia, N. Polymerase Chain Reaction. *Journal of Investigative Dermatology* **133**, 1–4 (2013).
89. Kubista, M. *et al.* The real-time polymerase chain reaction. *Molecular Aspects of Medicine* **27**, 95–125 (2006).
90. *The Polymerase Chain Reaction*. (Birkhäuser Boston, 1994). doi:10.1007/978-1-4612-0257-8.
91. *PCR protocols*. (Humana Press, 2003).
92. Barghouthi, S. A. A Universal Method for the Identification of Bacteria Based on General PCR Primers. *Indian Journal of Microbiology* **51**, 430–444 (2011).
93. PCR Basics - UK. https://www.thermofisher.com/uk/en/home/life-science/cloning/cloning-learning-center/invitrogen-school-of-molecular-biology/pcr-education/pcr-reagents-enzymes/pcr-basics.html.
94. Adams, G. A beginner's guide to RT-PCR, qPCR and RT-qPCR. *The Biochemist* **42**, 48–53 (2020).
95. Liu, H. Y. *et al.* Polymerase Chain Reaction and Its Application in the Diagnosis of Infectious Keratitis. *Med Hypothesis Discov Innov Ophthalmol* **8**, 152–155 (2019).
96. Doyle, T. *et al.* The importance of HIV RNA detection below 50 copies per mL in HIV-positive patients on antiretroviral therapy: an observational study. *The Lancet* **383**, S44 (2014).
97. Davenport, M. *et al.* New and developing diagnostic technologies for urinary tract infections. *Nat Rev Urol* **14**, 296–310 (2017).

98. Mach, K. E., Wong, P. K. & Liao, J. C. Biosensor diagnosis of urinary tract infections: a path to better treatment? *Trends in Pharmacological Sciences* **32**, 330–336 (2011).

99. Hudecova, I. Digital PCR analysis of circulating nucleic acids. *Clinical Biochemistry* **48**, 948–956 (2015).

100. Li, M., Diehl, F., Dressman, D., Vogelstein, B. & Kinzler, K. W. BEAMing up for detection and quantification of rare sequence variants. *Nat Methods* **3**, 95–97 (2006).

101. Dressman, D., Yan, H., Traverso, G., Kinzler, K. W. & Vogelstein, B. Transforming single DNA molecules into fluorescent magnetic particles for detection and enumeration of genetic variations. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 8817–8822 (2003).

102. Whitesides, G. M. The origins and the future of microfluidics. *Nature* **442**, 368–373 (2006).

103. Guo, F. *et al.* Valve-based microfluidic device for droplet on-demand operation and static assay. *Appl. Phys. Lett.* **97**, 233701 (2010).

104. Hindson, B. J. *et al.* High-Throughput Droplet Digital PCR System for Absolute Quantitation of DNA Copy Number. *Anal. Chem.* **83**, 8604–8610 (2011).

105. Notomi, T. Loop-mediated isothermal amplification of DNA. *Nucleic Acids Research* **28**, 63e–663 (2000).

106. Nagamine, K., Hase, T. & Notomi, T. Accelerated reaction by loop-mediated isothermal amplification using loop primers. *Molecular and Cellular Probes* **16**, 223–229 (2002).

107. Mori, Y. & Notomi, T. Loop-mediated isothermal amplification (LAMP): a rapid, accurate, and cost-effective diagnostic method for infectious diseases. *Journal of Infection and Chemotherapy* **15**, 62–69 (2009).

108. Ihira, M. *et al.* Rapid Diagnosis of Human Herpesvirus 6 Infection by a Novel DNA Amplification Method, Loop-Mediated Isothermal Amplification. *Journal of Clinical Microbiology* **42**, 140–145 (2004).

109. Notomi, T., Mori, Y., Tomita, N. & Kanda, H. Loop-mediated isothermal amplification (LAMP): principle, features, and future prospects. *J Microbiol.* **53**, 1–5 (2015).

110. Bruce, K. L., Leterme, S. C., Ellis, A. V. & Lenehan, C. E. Approaches for the detection of harmful algal blooms using oligonucleotide interactions. *Analytical and Bioanalytical Chemistry* **407**, 95–116 (2015).

111. Lakowicz, J. R. *Principles of Fluorescence Spectroscopy*. (Springer, 2006).

112. Fluorescence Fundamentals - UK. https://www.thermofisher.com/uk/en/home/references/molecular-probes-the-handbook/introduction-to-fluorescence-techniques.html.

113. Albani, J. R. *Structure and dynamics of macromolecules: absorption and fluorescence studies*. (Elsevier, 2004).

114. Dragan, A. I. *et al.* SYBR Green I: Fluorescence Properties and Interaction with DNA. *J Fluoresc* **22**, 1189–1199 (2012).

115. SYBR Green for qPCR | Bio-Rad. https://www.bio-rad.com/featured/en/sybr-green-for-qpcr.html.

116. SYBR® Green Based Quantitative PCR. *Sigma-Aldrich* https://www.sigmaaldrich.com/life-science/molecular-biology/pcr/quantitative-pcr/sybr-green-based-qpcr.html.

117. Zipper, H., Brunner, H., Bernhagen, J. & Vitzthum, F. Investigations on DNA intercalation and surface binding by SYBR Green I, its structure determination and methodological implications. *Nucleic Acids Res* **32**, e103 (2004).

118. Duprey, J.-L. H. A. *et al.* Rationalisation of a mechanism for sensing single point variants in target DNA using anthracene-tagged base discriminating probes. *Org. Biomol. Chem.* **16**, 6576–6585 (2018).

119. Farrell, R. E. *RNA methodologies: laboratory guide for isolation and characterization*. (Academic Press, an imprint of Elsevier, 2017).

120. Ducheyne, P. *Comprehensive biomaterials*. (Elsevier, 2011).

121. How TaqMan Assays Work - UK. //www.thermofisher.com/uk/en/home/life-science/pcr/real-time-pcr/real-time-pcr-learning-center/real-time-pcr-basics/how-taqman-assays-work.html.

122. PCR Protocol for Taq DNA Polymerase | NEB. https://international.neb.com/Protocols/0001/01/01/taq-dna-polymerase-with-standard-taq-buffer-m0273.

123. Cao, H. & Shockey, J. M. Comparison of TaqMan and SYBR Green qPCR Methods for Quantitative Gene Expression in Tung Tree Tissues. *J. Agric. Food Chem.* **60**, 12296–12303 (2012).

124. Nagy, A. *et al.* Evaluation of TaqMan qPCR System Integrating Two Identically Labelled Hydrolysis Probes in Single Assay. *Sci Rep* **7**, 41392 (2017).

125. Weinberger, K. M., Wiedenmann, E., Böhm, S. & Jilg, W. Sensitive and accurate quantitation of hepatitis B virus DNA using a kinetic fluorescence detection system (TaqMan PCR). *Journal of Virological Methods* **85**, 75–82 (2000).

126. Butler, J. M. *Advanced topics in forensic DNA typing: methodology*. (Elsevier/Academic Press, 2012).

127. Pacheco-Gómez, R. *et al.* Detection of Pathogenic Bacteria Using a Homogeneous Immunoassay Based on Shear Alignment of Virus Particles and Linear Dichroism. *Analytical Chemistry* **84**, 91–97 (2012).

128. Carr-Smith, J. *et al.* Polymerase Chain Reaction on a Viral Nanoparticle. *ACS Synthetic Biology* **4**, 1316–1325 (2015).

129. Ali, A. *et al.* Combining bacteriophage engineering and linear dichroism spectroscopy to produce a DNA hybridisation assay. *RSC Chem. Biol.* 10.1039.D0CB00135J (2020) doi:10.1039/D0CB00135J.

130. Kehoe, J. W. & Kay, B. K. Filamentous Phage Display in the New Millennium. *Chemical Reviews* **105**, 4056–4072 (2005).

131. Papavoine, C. H. M., Boukje, C. E. C., Rutger, F. H. A., Ruud, K. N. H. & Cornelis, H. W. Solution Structure of the M13 Major Coat Protein in Detergent Micelles: A Basis for a Model of Phage Assembly Involving Specific Residues. *Journal of Molecular Biology* **282**, 401–419 (1998).

132. Inovirus. https://viralzone.expasy.org/558?outline=all_by_species.

133. Vos, W. L., Nazarov, P. V., Koehorst, R. B. M., Spruijt, R. B. & Hemminga, M. A. From 'I' to 'L' and back again: the odyssey of membrane-bound M13 protein. *Trends in Biochemical Sciences* **34**, 249–255 (2009).

134. Stopar, D., Spruijt, R. B., Wolfs, C. J. A. M. & Hemminga, M. A. Protein–lipid interactions of bacteriophage M13 major coat protein. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1611**, 5–15 (2003).

135. Tridgett, M., Lloyd, J. R., Kennefick, J., Moore-Kelly, C. & Dafforn, T. R. Mutation of M13 Bacteriophage Major Coat Protein for Increased Conjugation to Exogenous Compounds. *Bioconjugate Chemistry* **29**, 1872–1875 (2018).

136. Tridgett, M. *et al.* Linear dichroism of visible-region chromophores using M13 bacteriophage as an alignment scaffold. *RSC Advances* **8**, 29535–29543 (2018).

137. Clack, B. A. & Gray, D. M. Flow linear dichroism spectra of four filamentous bacteriophages: DNA and coat protein contributions. *Biopolymers* **32**, 795–810 (1992).

138. Little, H. A. The Development of Novel Diagnostic Sensors Based on Linear Dichroism Spectroscopy. (University of Birmingham, 2016).

139. Ali, A. The Control and Detection of Biomolecules Using Functional DNA. (University of Birmingham, 2018).

Chapter 2: Techniques

2.1 Introduction

This chapter serves to introduce the main analytical techniques used throughout this work. This includes a description of a range of optical spectroscopy techniques including ultraviolet-visible spectroscopy (UV-Vis), fluorescence spectroscopy, circular dichroism (CD) and linear dichroism (LD). In addition to these there will be an introduction to techniques used for the analysis of oligonucleotides; high performance liquid chromatography (HPLC) and mass spectrometry (MS).

2.2 Absorbance Spectroscopy

The work in this thesis relies on a range of absorbance spectroscopy techniques including UV-Vis, CD and LD, below the theory behind these techniques will be discussed.

The majority of spectroscopic phenomena are a result of an interaction between a molecule and photons. The molecule can either absorb or scatter the interacting photon.[1] The simplest form of spectroscopy is absorbance spectroscopy, which measures the quantity of light absorbed by the molecules at a given wavelength.[1] When light passes through the compound, energy from the photons are used to excite an electron from the ground state, highest occupied molecular orbital (HOMO) into an excited state, the lowest unoccupied molecular orbital (LUMO).[2,3] Photons of different wavelengths possess different energies which can be calculated using **Equation 2.1**, where ΔE is the energy, h is Planck's constant, c is the speed of light and λ is the wavelength.[4]

$$\Delta \mathrm{E} = \frac{hc}{\lambda}$$

***Equation 2.1**: The equation for determining the energy gap between the bonding and antibonding orbital of a molecule based on the absorption of a photon of light. **ΔE** is the energy gap between the bonding and antibonding orbitals, **h** is Planck's constant, **c** is the speed of light and **λ** is the wavelength of the indecent photon.*[4,5]

Smaller energy gaps such as those involved in n – π* (non-bonding orbital to anti π bonding orbital) transitions require less energy, and therefore a longer wavelength of light. Larger energy gaps such as π – π* (π bonding orbital to anti π bonding orbital) transitions on the other hand require more energy, a higher frequency and as a result a shorter wavelength of light. Meanwhile transitions involving σ or σ* orbitals all possess much larger energy gaps, subsequently higher energy photons and therefore all require wavelengths <190 nm (**Figure 2.1**).[1,3]

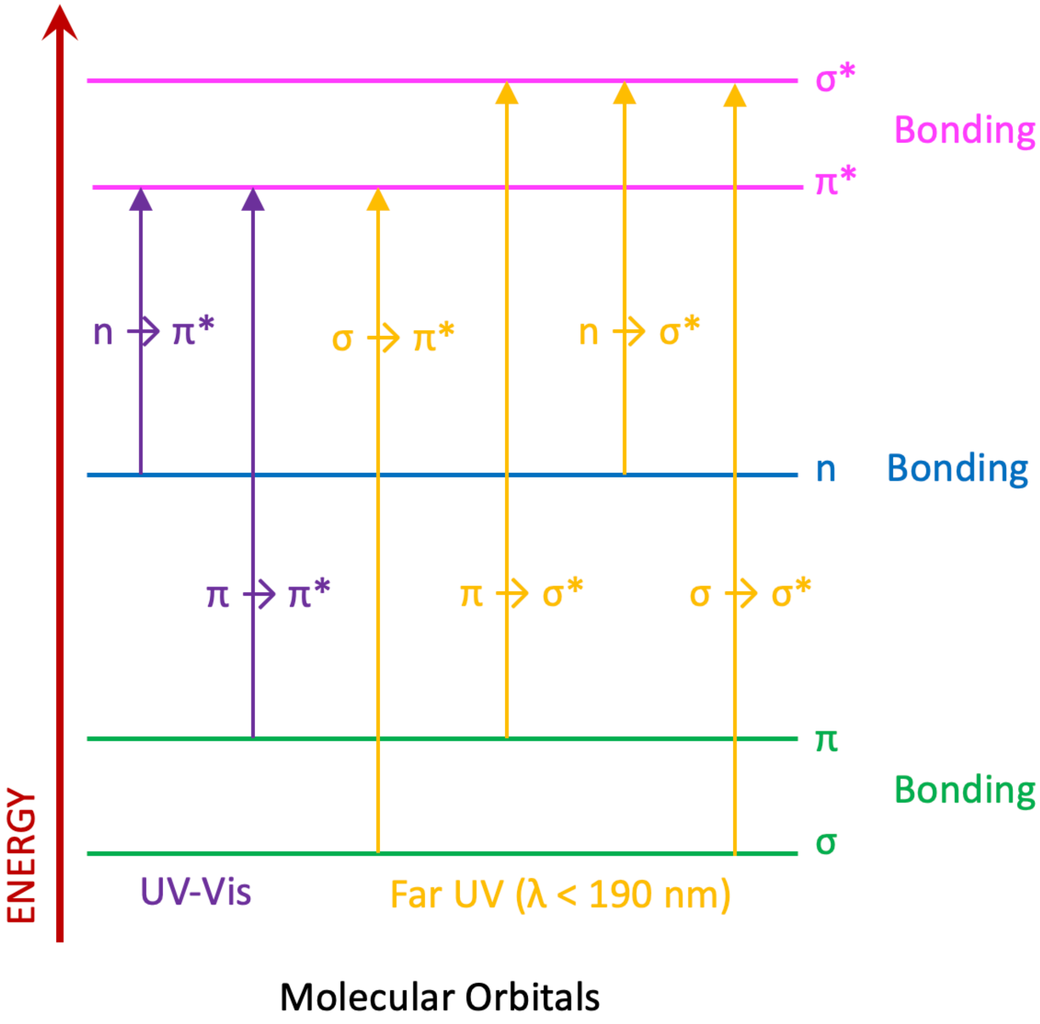


Figure 2.1*: The transitions involved in absorbing species containing π, σ and n electrons. Blue arrows can be achieved with light ranging between UV and Visible. Orange arrows can only be achieved with wavelengths shorter than 190 nm. Image adapted from Bonardi et al*[6]

As a result of the different energy gaps different wavelengths will be absorbed by a molecule, thus there will be less photons of a particular wavelength exiting a sample compared to what entered. A spectrophotometer is capable of measuring this change thereby enabling a quantitative measure of the intensity of photons at every wavelength using **Equation 2.2**. This gives rise to an absorbance spectrum which is a plot of absorption versus wavelength giving information about the molecule being measured.[1,2,7–9]

$$A = \log_{10}\left(\frac{I_0}{I}\right)$$

***Equation 2.2**: The equation for calculating absorption at a given wavelength. **A** is absorption, **I_0** is the initial intensity of light, and **I** is the transmitted intensity of light.*

Interestingly; if we visualise absorption as the electromagnetic field of a photon pushing electron density from the ground state to the excited state, we can consider the direction of the push as the polarisation of transition, and the force of the push is the intensity.[1,2,9] Together the polarisation and intensity of a transition is known as the electric transition dipole moment (**μ**), and each transition can be thought of as an unique light harvesting 'antenna'.[1] When an 'antenna' is aligned fully parallel to an incident of light the maximum amount of absorbance is possible, whereas when the 'antenna' is aligned fully perpendicular, no light can be absorbed.[1] In isotropic samples this has little to no effect on the absorbance spectra, however, if a sample is aligned it can lead to a determination of polarisation.[1,10–12]

2.2.1 Ultraviolet Visible Spectroscopy

The most used method of absorbance spectroscopy is UV-Vis Spectroscopy. This technique can be used to determine the presence and concentration of a particular substance. Isotropic light is used to irradiate a sample and photons with the appropriate energy are absorbed, whereas the rest are transmitted.[3,13,14] The wavelength that absorbs the most light is referred to as the lambda max (λ_{max}) which is proportional to the concentration of a moieties. The concentration can then be calculated using the Beer-Lambert Law (**Equation 2.3**).[1–3,8]

$$A = \varepsilon C l$$

***Equation 2.3**: The Beer-Lambert Law, **A** is absorbance at the λ_{max} wavelength, **ε** is the extinction coefficient of the molecule, **C** is the concentration of the sample and **l** is the path length of the sample.*[1,2]

2.2.2 Circularly Polarised Light

An electromagnetic wave is composed of electric and magnetic fields orthogonal to one another.[15] A polarisation of light only has the electric field in one plane. Multiple waves of light are able to interact with one another, the classical example of this is constructive and destructive interference. Constructive interference is when two waves of the same polarisation are in phase, they have the same wavelength and their maximum point and minimum point are the same, they combine to produce a wave of larger amplitude.[16,17] Destructive interference on the other hand is when two waves of the same polarisation are out of phase, they have the same wavelength however, their maximum point and minimum point are the opposite, they combine to produce a wave of smaller amplitude (**Figure 2.2**).[16,17]

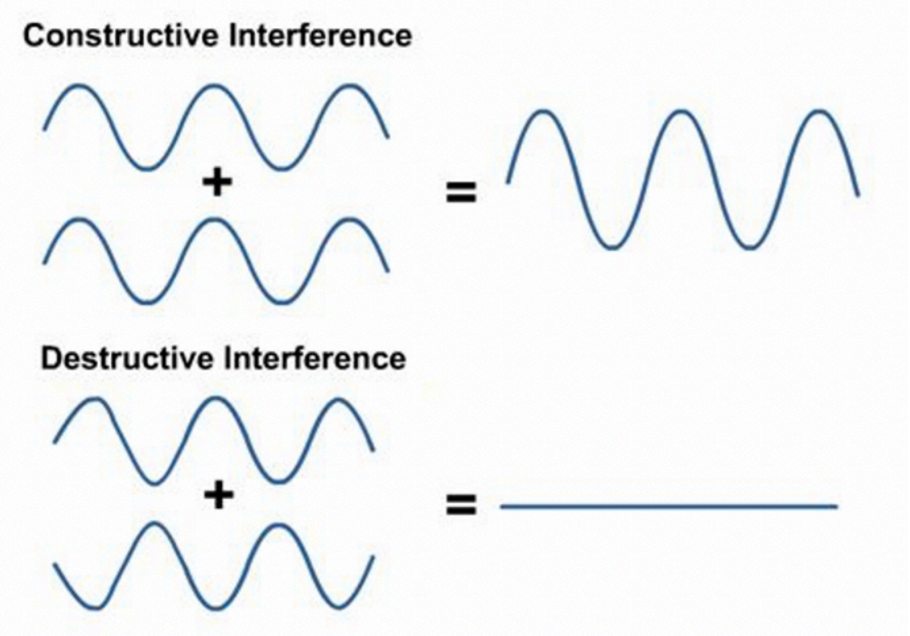


Figure 2.2*: (Top) Two waves in phase, which constructively combine to form a wave of larger amplitude. (Bottom) Two waves out of phase, which destructively combine to produce a wave of smaller amplitude. Image taken from Gitam school of physics.*[16]

In a system with two waves of equal magnitude, propagating in the same direction, yet out of phase by π/2 (a quarter of a wavelength) an interesting phenomenon occurs. The wave being out of phase by a quarter of a wavelength results in the light being circularly polarised (**Figure 2.3**). A CD spectropolarimeter produces circularly polarised light using a photoelastic modulator (PEM), a piece of crystalline quartz coupled to a piece of isotropic quartz. Upon the application of an external alternative current (AC) voltage source, the PEM converts a single linear polarisation of light into two linear polarisations orthogonal to each other and out of phase by a quarter of a wavelength.[1]

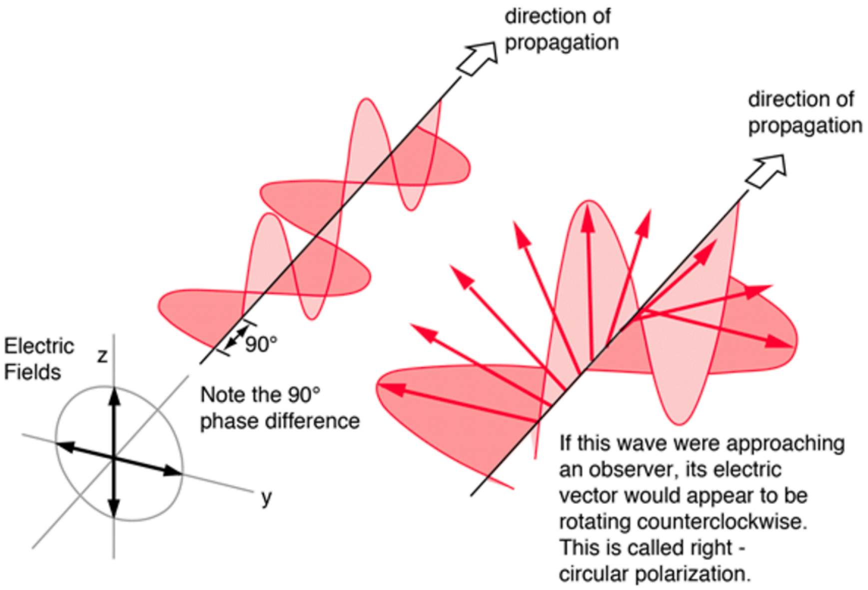


Figure 2.3*: A diagram showing the generation of circularly polarised light, two polarisations of light orthogonal to one another and out of phase by a quarter of a wavelength, appears to be rotating in a circular motion. Image taken from HyperPhysics.*[18]

This circular polarisation of light enables for the study of chiral molecules using an absorbance technique called CD.

2.2.2.1 Circular Dichroism

The majority of biological molecules, such as DNA and proteins, possess some degree of chirality or helicity both as a result of their subunits being chiral and their folding to form secondary and tertiary structures creating helices.[19] CD is a differential absorbance technique that enables the study of samples with chirality and for probing the higher order structures of macromolecules. Using **Equation 2.4** CD can be calculated by measuring the difference in absorbance of left handed (A_l) and right handed (A_r) light.[1,12]

$$CD = A_l - A_r$$

***Equation 2.4**: The equation for calculating the differential CD signal at a given wavelength. **CD** is the circular dichroism signal at a given wavelength, **A_l** is the absorbance of left-handed circularly polarised light and **A_r** is the absorbance of right-handed circularly polarised light.*

When a molecule absorbs one polarisation of light more than the other, a CD signal is observed. As a result the CD signal can be either positive or negative, with positive CD signals being a result of a greater absorbance of left-handed circularly polarised light and negative signals a greater absorbance of right-handed circularly polarised light.[1,20,21] This technique makes CD and invaluable tool for the determination of the secondary structure of proteins and DNA amongst other things.

2.2.3 Linearly Polarised Light

A modification to a CD spectropolarimeter, results in the circularly polarised light produced being converted into two linear polarisations of light orthogonal to one another.[1]

2.2.3.1 Linear Dichroism

Linear dichroism enables the identification of different orientation parameters of molecules. This technique measures the difference in absorptions of two linear polarisations of light, 90° to one another.[22] The difference between the absorption of parallel light ($A_{\parallel}$) and perpendicular light ($A_{\perp}$) to a defined orientation (*z*-axis) gives rise to an LD signal at a given wavelength, (**Equation 2.5**) thus aiding in the determination of the orientation of the **μ** of a molecule.[1]

$$LD = A_{\parallel} - A_{\perp}$$

Equation 2.5**: The equation for calculating the differential LD signal at a given wavelength.* ***LD *is the linear dichroism signal at a given wavelength,* ***A||*** *is the absorbance of parallelly polarised light and* ***A⊥*** *is the absorbance of perpendicularly polarised light.*

Malus' law (**Equation 2.6**) states that "the intensity of a beam of plane-polarized light after passing through a rotatable polarizer varies as the square of the cosine of the angle through which the polarizer is rotated from the position that gives maximum intensity."[1,23–26]

$$I = I_0 cos^2\theta$$

Equation 2.6**: The equation for calculating the intensity of transmitted light after passing through a polariser at different angles.* ***I *is the intensity of transmitted light,* ***I₀*** *is the initial intensity of light and* ***Θ*** *is the angle which the polariser has rotated from the plane of maximum intensity.*

Based on Malus' Law when two perfect polariser are aligned the intensity of transmitted light will be at a maximum and when they are at 90° to one another the intensity of transmitted light will be 0 with a varying quantity between the two (**Figure 2.4**).[1,24,26,27]

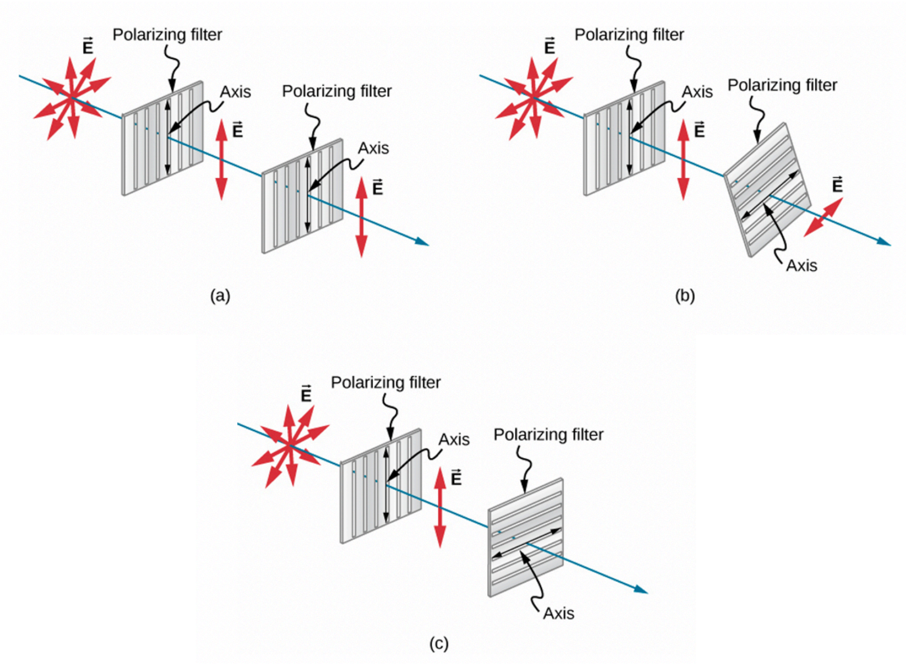


***Figure 2.4**: A diagram showing the effect of polarising filters, in all diagrams the first filter polarises the light. (a) All the polarised light passes through the second filter as the axis is parallel to the first. (b) Only part of the light is transmitted, as the second filter is rotated, the quantity of transmitted light is calculated based on **Equation 2.6**. (c) None of the light is able to be transmitted as the second filter is perpendicular to the first..*[23,24,26,27] *Image adapted from Zanelli*[28]

In an LD system we can consider the polariser inside the spectropolarimeter to be polariser 1 and the molecule of interest to be polariser 2, however, instead of transmitting light, it is absorbed.[1] Therefore if the **μ** of a molecule is aligned completely parallel to the orientation direction only a parallelly polarised photon can be absorbed (**Equation 2.7**).[1]

$$LD = A_{\parallel} > 0$$

***Equation 2.7**: The equation for LD signal when light is absorbed in the parallel direction.*

Conversely, when polariser 1 is rotated 90° to be perpendicular to the alignment direction, if the **μ** of a molecule is aligned completely perpendicular to the orientation direction only a perpendicularly polarised photon can be absorbed (**Equation 2.8**).[1]

$$LD = -A_{\perp} < 0$$

Equation 2.8: *The equation for LD signal when light is absorbed in the perpendicular direction.*

More commonly though, the **μ** sits somewhere between the two extremes. As a result, the absorption intensity and therefore LD varies, in concordance with **Equation 2.9**. Where S is the orientation parameter, which can be taken to be a value between 1 (a perfectly aligned sample) and 0 (a completely isotropic sample), A_{iso} is the absorbance of an unaligned (completely isotropic) sample and α is the angle between **μ** and the laboratory defined parallel axis. When α is the 'magic angle' of 54.7°, $3\cos^2\alpha$ is equal to 1, and as a result, the LD signal will always be 0 at this point.[1]

$$LD = \frac{3}{2}\, S\, A_{iso}(3cos^2\alpha - 1)$$

Equation 2.9: *The equation for calculating LD signal.* ***LD*** *is the linear dichroism signal,* ***S*** *is the orientation parameter,* ***A_{iso}*** *is the absorbance of an isotropic sample and* ***α*** *is the angle between the electric transition dipole moment and the laboratory defined z-axis.*

While differential LD is an extremely sensitive technique requiring only small concentrations of sample, in order for quantitative analysis another measurement is required, isotropic absorbance (A_{iso}) of an unoriented sample. With A_{iso} **Equation 2.9** can be rewritten as reduced LD (LD^r) (**Equation 2.10**). LD^r is independent of path length and concentration, therefore the magnitude of the LD^r signal is a response of the orientation of the **μ** alone.[1] When the orientation of the molecule itself is known, LD^r can be used to calculated the value of α and this the polarisation of the electric dipole transition moment.[1]

$$LD^r = \frac{LD}{A_{iso}} = \frac{3}{2}S(3cos^2\alpha - 1)$$

***Equation 2.10**: The equation for LD^r signal at a given wavelength. **LD^r** is the reduced linear dichroism signal, **LD** is the linear dichroism signal, **A_{iso}** is the absorbance of an isotropic sample, **S** is the orientation parameter and **α** is the angle between the electric transition dipole moment and the laboratory defined parallel axis.*

2.2.3.2 Alignment of Samples for use in LD

In order to measure LD samples must first be aligned. The physical properties of the molecule are the key in determining which orientation methodology is used (**Table 2.1**).

***Table 2.1**: The molecular orientation methods for LD.*

Alignment Technique	Example Molecule
Couette Flow	M13 Bacteriophage[29–31]
Stretched Film	Anthracene[32]
Gel Deformation	Membrane Vesicles[11]
Electric Field	Chloroplasts[33]
Crystallisation	Proteins[34]
Liquid Crystal	Naphthalene derivatives[35]

While there are a multitude of ways to align molecules the most common methods will be discussed below.

2.2.3.2.1 Shear Flow Orientation

Molecules which possess a high aspect ratio (long and thin) can be solvated and flowed past a stationary surface to generate shear forces. The shear forces cause the long axis of the molecule to align along the flow direction. The flow rate and resultant shear forces can be generated by a flow system such as a HPLC pump,[36] however the requirement of a large sample and the possible introduction of air bubbles, hinder the usefulness of this technique.[1] These problems are alleviated through the use of a Couette cell (**Figure 2.5**), it is worth noting that this is often misnamed as cuvette, but the name originates from its inventor Maurice Couette.[1,37–39] A Couette cell is composed of an outer rotating quartz capillary and a stationary inner quartz rod, with the sample occupying the void between the two.[22,29,30,37] As the outer capillary rotates, the shear forces are produced aligning molecules in solution.[1,40–43] The development of microvolume Couette cells enabled microlitre samples to be analysed and is common place in an LD lab.[39,40,44–46] In a micro Couette system, the alignment of the molecular long axis is considered to be parallel to the floor (*z* axis) and the parallel polarisation of light.[22,44,47]

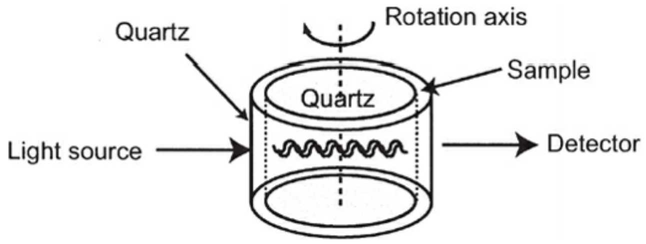


***Figure 2.5:** Diagrammatic depiction of Couette flow alignment for linear dichroism. Image adapted from Nordén et al.*[1]

2.2.3.2.2 Stretched Film Orientation

Small molecules cannot be aligned through shear flow, due to an insufficient level of force being exerted on the molecule. Through the adsorption of molecules onto a polymer film which is mechanically stretched, alignment occurs (**Figure 2.6**). The stretching aligns the long axis of the molecule parallel to the stretch direction.[22,48,49] Polyethylene (PE) possesses a microcrystalline structure and as a result produces a molecular orientating environment when stretched.[50] When the PE is stretched the crystalline structure forms uniform grooves in the same direction, these grooves can be visualised as a field after it has been ploughed, the adsorbed molecules are then directed to align with their long axis in the groove.[32,51] The analyte is dissolved in a volatile solvent, and deposited onto the surface of the stretched film, the solvent is allowed to evaporate and the LD measurement is recorded. PE is suitable for aligning only non-polar molecules meanwhile oxidised polyethylene (PE^{OX}) films possess are able to align both polar and non-polar molecules, thus making them the favourable choice for small molecule alignment.[32,49,51] The stretch direction, in a standard system, can then be defined to be parallel to the floor (z axis) and thus the parallel polarisation of light.[32,51] This alignment results in the molecules being fully aligned, consequently reduced linear dichroism (LD^r) can be used to determine the orientation of the electric transition dipole moment ($\boldsymbol{\mu}$).[49]

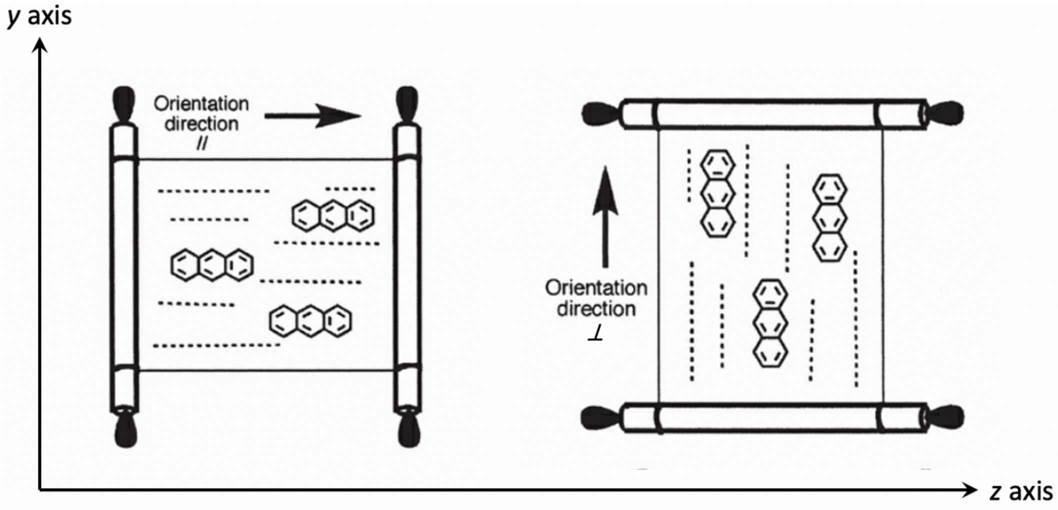


Figure 2.6: *Diagrammatic depiction of stretched film linear dichroism. Anthracene adsorbed onto the film aligns along the long axis and shows how the orientation of the stretch direction along either the z or y axis can change the absorbance direction. Image adapted from Nordén et al.*[1]

2.2.3.3 Linear Dichroism of Bacteriophage

Filamentous bacteriophage such as M13 bacteriophage possess a high aspect ratio, being 900 nm long and 7 nm in diameter,[52] this makes it an ideal target for LD (**Figure 2.7**). Filamentous bacteriophages, such as M13, are LD active and possesses a positive LD signal between 270 nm and 310 nm as a result of the aromatic amino acids in pVIII protein aligning parallel to the bacteriophages long axis, a negative LD signal between 240 nm and 270 nm is a result of the DNA bases aligning perpendicular to the long axis. Below 240 nm there is a positive signal near 224 nm which was suggested to be a result of tyrosine and tryptophan aligning parallel to the long axis. The final positive peak at 206 nm is a result of $\pi - \pi^*$ transitions in the α-helix of the major coat proteins (**Figure 2.7**).[31]

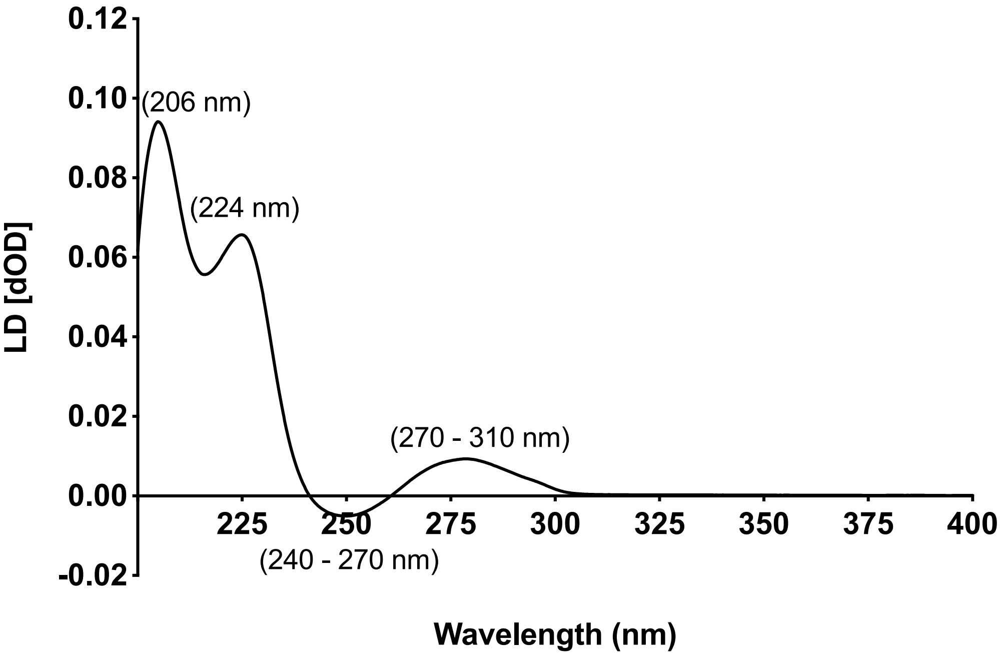


Figure 2.7*: An LD spectrum of M13 bacteriophage, with the transition wavelengths labelled.*

2.3 Fluorescence Spectroscopy

Absorbance spectroscopy is not the only identification technique employed over the course of this work. Through the incorporation of an intercalating dye the amplification of DNA can be monitored using fluorescence spectroscopy below the theory behind fluorescence spectroscopy will be discussed. We previously introduced the concept of energy gaps between molecular orbitals, and for the purpose of absorbance spectroscopy this is sufficient, however fluorescence spectroscopy called for a further understanding of energy levels.[53] It is possible to visualise energy levels in a molecule using a Jablonski diagram (**Figure 2.8**) which represents the different energy levels of a molecule. The baseline represents the ground state of an electron whereas the upper bold lines represent different excited electronic states, S_1 and S_2 representing singlet states, where the overall spin is 0, while T_1 and T_2 represent triplet states where the overall spin is 1.[54] In every state, be that ground or excited there are multiple vibrational levels, which are represented by the thinner lines above the electronic states.[55,56] When a photon of sufficient energy incidents on a molecule, an electron is raised from a HOMO to a LUMO, however it is not as simple as going from S_0 to $S_{1;}$ it is more than likely the electron will be raised to a vibrational level within the higher electronic state. This electron will then return to ground state through a combination of steps with the favoured method being the one that returns the electron to the ground state as quickly as possible.[53,55,57] The most common relaxation step, seen during almost all excitations is vibrational relaxation. Collisions of the excited molecule with other molecules will lead to a loss of energy, thus causing an electron to drop from a high vibrational level to the next lowest electronic level; this is the fastest method of relaxation. Internal conversions (IC) is an intermolecular process of nonradiative relaxation (no photon is emitted) between two electronic states of the same multiplicity.[57] IC usually occurs when two electronic energy levels are close enough for two vibrational levels to overlap, thus causing the electron to transition to a lower energy state before vibrational relaxation. Due to the energy gap between the ground state and first state being larger than between any other, de-excitation from S_1 to S_0 is far less efficient than relaxation to the first state, however it does occur in aliphatic compounds.[57] Therefore the more favourable method of de-excitation for aromatic compounds is a radiative transition, through the emission of a photon. [55,57,58]

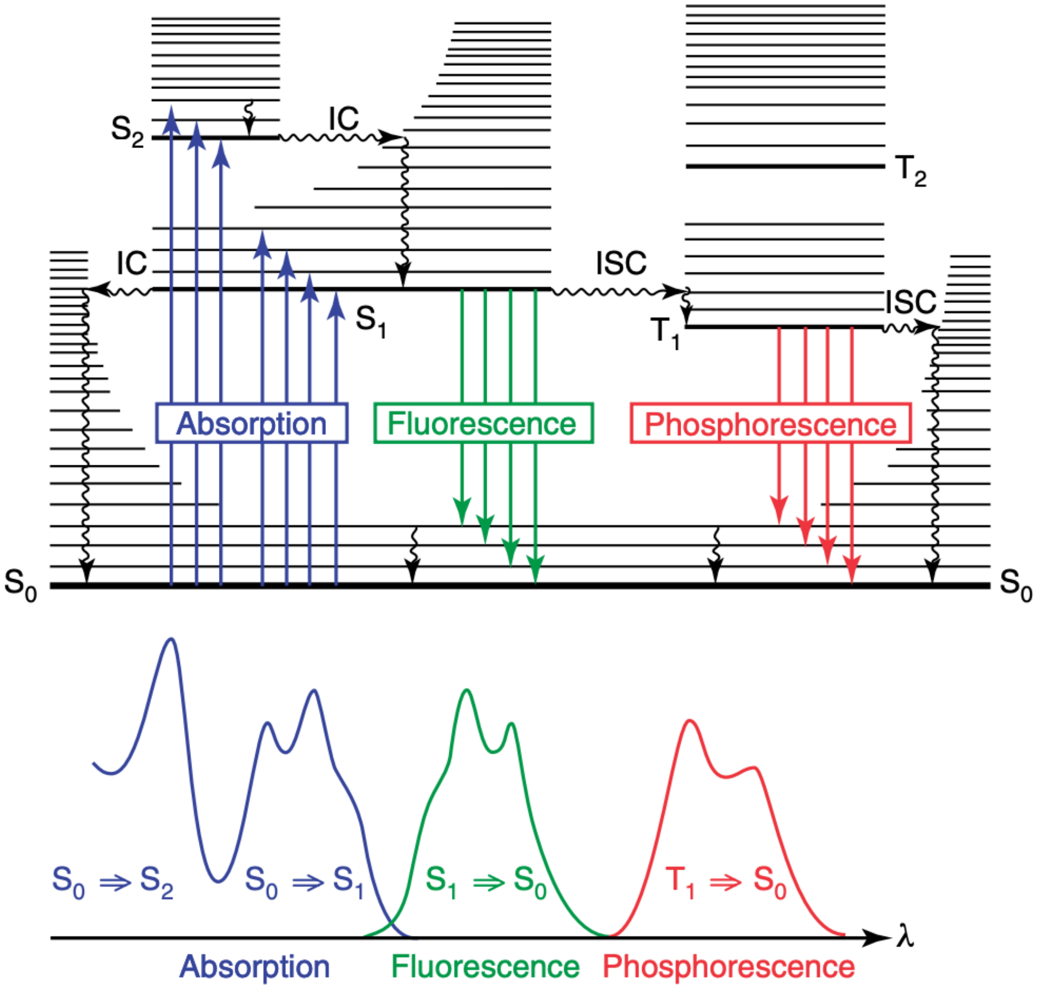


***Figure 2.8**: Jablonski diagram and relative positions of the spectra. S_0 Ground state. S_1 First singlet electronic state, S_2 Second singlet electronic state. T_1 First triplet electronic state. T_2 Second triplet electronic state. IC internal conversions. ISC Intersystem crossing. Image taken from Valeur et al.[57]*

While IC is an intermolecular process of nonradiative relaxation, between two electronic states of the same multiplicity it is also possible for a similar process to occur between electronic states of different multiplicities; this process is known as intersystem crossing (ISC).[57] Following ISC the electron will vibrationally relax to the lowest energy Triplet state, T_1.[53] The ground state of a molecule, S_0, is a singlet state, therefore in order for T_1 to return either another ISC to a vibrational level in S_0 of a similar energy or a radiative process must occur. This radiative process is called phosphorescence and occurs when the electron transitions from a triplet to singlet state by emitting a photon.[55–58] It is however, far more common for emission to occur without ISC. The de-excitation of from S_1 to S_0 through the emission of a photon is known as fluorescence. The photons emitted from fluorescence are of a lower energy than those absorbed by the molecule as a result various relaxation steps, with the gap between the absorbance λ_{max} and emission λ_{max} being known as the Stokes shift.[59] Fluorescence of a molecule will always occur from the lowest vibrational level of S_1 to the ground state, the wavelength of the photon emitted will correspond to the energy gap of a $\pi^* - \pi$ or $\pi^* - n$ transition depending on which is smaller, as a result the emission spectra of a particular molecule will always be the same with only a variation in magnitude depending on excitation energy.[57,58] As the trace of emission spectra will always be the same, it is possible to measure the fluorescence of a molecule using a spectrofluorometer. The property of molecules to produce a unique emission spectrum regardless of excitation wavelength enables for the monitoring of an increase in fluorescence under varying conditions. One such condition such as that used in this work is the increase in fluorescence intensity of an intercalating dyes in the presence of duplex DNA.[60–64]

2.4 High Performance Liquid Chromatography

In complex reaction mixtures, such as an DNA amplification solution after heating, it can be necessary to separate the constituents in order to assess individual products. Chromatography, such as that used in purification of synthetic products, is a separation technique based upon relative polarities.[65] While column chromatography relies primarily on gravity, HPLC operates using high pressure to separate the mixture.[66,67] Aside from the increased pressure HPLC functions in a similar manner to column chromatography: a column is filled with a stationary, polar phase, usually silica and a non-polar mobile phase is flowed through the column. When a mixture is introduced onto the column the chemical structure will determine the rate in which is flows through, with non-polar molecules preferring to interact with the mobile phase and therefore eluting first, whereas more polar molecules will interact with the stationary phase and as a result take longer to travel through the column.[66–68] HPLC offers the ability to monitor the spectroscopic properties (absorbance/fluorescence) of the eluent to determine the purity of the molecules of interest.

2.4.1 Reversed-Phase High Performance Liquid Chromatography

While HPLC is an extremely useful technique for the separation of mixtures, the majority of biological molecules are highly polar, therefore unsuitable for regular HPLC. Reversed-phase HPLC (RP-HPLC) provides a solution to this by replacing the polar stationary phase of with a non-polar hydrocarbon coated silica particles[69] and substituting the non-polar mobile phase for one of a polar nature (water/methanol/acetonitrile). As a result, the retention times of the molecules are reversed relative to normal phase HPLC.[70,71] Polar molecules such as DNA and proteins elute first whereas non-polar impurities are retained on the stationary phase.[71] RP-HPLC was therefore used as an efficient and simple purification process for the biological mixtures reported on in this thesis.

2.5 Mass Spectrometry

As with HPLC, if a mixture can be separated the individual components can be analysed, the complex structure of DNA means an idea method to differentiate between sequences is using the compounds mass. MS is a technique used to determine the mass of a compound. A mass spectrometer can be divided into four sections, sample injection, ion generation, mass separation, and ion detection.[2,72] Samples are analysed according to their mass to charge ratio, as a result, prior to analysis they must ionised. The ionisation technique depends on the nature of the compound, with the samples reported on in this thesis being ionised using electrospray ionisation (ESI).[2,72,73] ESI is 'soft' ionisation technique that employs the use of an electrified needle to simultaneously ionise and inject the sample as an aerosol. It is worth noting that ESI forms multiple charged species, resulting in a charge greater than 1 in molecules such as DNA (e.g. $[DNA]^{1+}$ and $[DNA]^{2+}$ etc).[2,72,73] Once the molecules have been ionised, they are accelerated through the use a magnetic field. The magnetic field separates the ionised molecules based upon their mass to charge ratio, with lighter ions deflecting more drastically than their heavier counterparts.[2,73] This results in the various ions being separated and subsequently detected as distinct signals. The largest peak displayed on the graph is the molecular ion peak which can be used to determine the molecular weight of the compound.[2]

2.6 References

1. Nordén, B., Rodger, A. & Dafforn, T. *Linear dichroism and circular dichroism: a textbook on polarized-light spectroscopy*. (Royal Society of Chemistry, 2010).
2. Hesse, M., Meier, H. & Zeeh, B. *Spectroscopic Methods in Organic Chemistry*. (Thieme, 2008).
3. UV-visible absorption spectra. https://www.chemguide.co.uk/analysis/uvvisible/theory.html.
4. Flowers, P., Therpold, K. & Langley, R. 6.2 The Bohr Model. in *Chemistry* (2016).
5. Waerden, B. L. van der. *Sources of quantum mechanics*. (Dover, 2007).
6. Bonardi, A.-H., Dumur, F., Noirbent, G., Lalevée, J. & Gigmes, D. Organometallic vs organic photoredox catalysts for photocuring reactions in the visible region. *Beilstein J. Org. Chem.* **14**, 3025–3046 (2018).
7. Atkins, P. W. & Friedman, R. S. *Molecular Quantum Mechanics*. (Oxford University Press, 2010).
8. Hollas, J. M. *Modern spectroscopy*. (J. Wiley, 2004).
9. Atkins, P. W., De Paula, J. & Keeler, J. *Atkins' Physical chemistry*. (Oxford University Press, 2018).
10. Albinsson, B., Kubista, M., Norden, B. & Thulstrup, E. W. Near-ultraviolet electronic transitions of the tryptophan chromophore: linear dichroism, fluorescence anisotropy, and magnetic circular dichroism spectra of some indole derivatives. *J. Phys. Chem.* **93**, 6646–6654 (1989).
11. Kiss, L. I., Ganago, A. O. & Garab, Gy. I. Quantitative method for studying orientation of transition dipoles in membrane vesicles of spherical symmetry. *Journal of Biochemical and Biophysical Methods* **11**, 213–225 (1985).
12. Nordén, B. Applications of linear Dichroism Spectroscopy. *Applied Spectroscopy Reviews* **14**, 157–248 (1978).
13. Physics of Absorption • Ocean Optics Web Book. http://www.oceanopticsbook.info/view/absorption/physics_of_absorption.
14. Bielecki, Z. *et al.* Selected optoelectronic sensors in medical applications. *Opto-Electronics Review* **26**, 122–133 (2018).

15. Anatomy of an Electromagnetic Wave | Science Mission Directorate. https://science.nasa.gov/ems/02_anatomy.
16. Interference. *Gitam School Of Physics* https://gitamschoolofphysics.wordpress.com/interference/ (2016).
17. Constructive and Destructive Interference. https://www.phys.uconn.edu/~gibson/Notes/Section5_2/Sec5_2.htm.
18. Classification of Polarization. http://hyperphysics.phy-astr.gsu.edu/hbase/phyopt/polclas.html#c1.
19. Blackmond, D. G. The Origin of Biological Homochirality. *Cold Spring Harbor Perspectives in Biology* **2**, a002147–a002147 (2010).
20. Bulheller, B. M., Rodger, A. & Hirst, J. D. Circular and linear dichroism of proteins. *Physical Chemistry Chemical Physics* **9**, 2020 (2007).
21. Greenfield, N. J. Using circular dichroism spectra to estimate protein secondary structure. *Nat Protoc* **1**, 2876–2890 (2006).
22. Hicks, M. R., Kowałski, J. & Rodger, A. LD spectroscopy of natural and synthetic biomaterials. *Chemical Society Reviews* **39**, 3380 (2010).
23. Kahr, B. & Claborn, K. The Lives of Malus and His Bicentennial Law. *ChemPhysChem* **9**, 43–58 (2008).
24. Malus's Law. https://spie.org/publications/fg05_p03_maluss_law.
25. Volobuev, A. N. *Interaction of the electromagnetic field with substance*. (Novinka, 2013).
26. 1.8: Polarization. *Physics LibreTexts* https://phys.libretexts.org/Bookshelves/University_Physics/Book%3A_University_Physics_(OpenStax)/Map%3A_University_Physics_III_-_Optics_and_Modern_Physics_(OpenStax)/01%3A_The_Nature_of_Light/1.08%3A_Polarization (2016).

27. Nurushev, S. B., Runtso, M. F. & Strikhanov, M. N. *Introduction to polarization physics*. (Springer, 2013).
28. Zanelli, C. I. *et al.* Total non-elastic cross sections of neutrons on C, O, Ca, and Fe at 40.3 and 50.4 MeV. *Phys. Rev. C* **23**, 1015–1022 (1981).
29. Carr-Smith, J. *et al.* Polymerase Chain Reaction on a Viral Nanoparticle. *ACS Synthetic Biology* **4**, 1316–1325 (2015).
30. Pacheco-Gómez, R. *et al.* Detection of Pathogenic Bacteria Using a Homogeneous Immunoassay Based on Shear Alignment of Virus Particles and Linear Dichroism. *Analytical Chemistry* **84**, 91–97 (2012).
31. Clack, B. A. & Gray, D. M. Flow linear dichroism spectra of four filamentous bacteriophages: DNA and coat protein contributions. *Biopolymers* **32**, 795–810 (1992).
32. Razmkhah, K., Little, H., Sandhu, S., Dafforn, T. R. & Rodger, A. Optical properties of xanthene based fluorescent dyes studied by stretched-film linear dichroism. *RSC Adv.* **4**, 37510–37515 (2014).
33. Gagliano, A. G., Geacintov, N. E. & Breton, J. Orientation and linear dichroism of chloroplasts and sub-chloroplast fragments oriented in an electric field. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **461**, 460–474 (1977).
34. Allen, J. P. & Feher, G. Crystallization of reaction center from Rhodopseudomonas sphaeroides: preliminary characterization. *Proceedings of the National Academy of Sciences* **81**, 4795–4799 (1984).
35. Bauman, D. & Kuball, H.-G. UV-visible linear dichroism of naphthalene bicarboxylic acid derivatives dissolved in nematic liquid crystal. *Chemical Physics* **176**, 221–231 (1993).
36. Moore-Kelly, C., Welsh, J., Rodger, A., Dafforn, T. R. & Thomas, O. R. T. Automated High-Throughput Capillary Circular Dichroism and Intrinsic Fluorescence Spectroscopy for Rapid Determination of Protein Structure. *Anal. Chem.* **91**, 13794–13802 (2019).

37. Simonson, T. & Kubista, M. DNA orientation in shear flow. *Biopolymers* **33**, 1225–1235 (1993).
38. Nordh, J., Deinum, J. & Nordén, B. Flow orientation of brain microtubules studies by linear dichroism. *Eur Biophys J* **14**, (1986).
39. Marrington, R., Seymour, M. & Rodger, A. A new method for fibrous protein analysis illustrated by application to tubulin microtubule polymerisation and depolymerisation. *Chirality* **18**, 680–690 (2006).
40. Taylor, G. I. VIII. Stability of a viscous liquid contained between two rotating cylinders. *Phil. Trans. R. Soc. Lond. A* **223**, 289–343 (1923).
41. van Gorp, J. J., Vekemans, J. A. J. M. & Meijer, E. W. Facile synthesis of a chiral polymeric helix; folding by intramolecular hydrogen bonding. *Chem. Commun.* 60 (2004) doi:10.1039/b312407j.
42. *Modern techniques for circular dichroism and synchrotron radiation circular dichroism spectroscopy*. (IOS Press, 2009).
43. Bassindale, A. R., Codina-Barrios, A., Frascione, N. & Taylor, P. G. An improved phage display methodology for inorganic nanoparticle fabrication. *Chemical Communications* 2956 (2007) doi:10.1039/b702650a.
44. Marrington, R. *et al.* Validation of new microvolume Couette flow linear dichroism cells. *The Analyst* **130**, 1608 (2005).
45. Marrington, R., Dafforn, T. R., Halsall, D. J. & Rodger, A. Micro-Volume Couette Flow Sample Orientation for Absorbance and Fluorescence Linear Dichroism. *Biophysical Journal* **87**, 2002–2012 (2004).
46. Rodger, A. *et al.* Looking at long molecules in solution: what happens when they are subjected to Couette flow? *Physical Chemistry Chemical Physics* **8**, 3161 (2006).
47. Dafforn, T. R., Rajendra, J., Halsall, D. J., Serpell, L. C. & Rodger, A. Protein Fiber Linear Dichroism for Structure Determination and Kinetics in a Low-Volume, Low-Wavelength Couette Flow Cell. *Biophysical Journal* **86**, 404–410 (2004).

48. Rodger, A., Dorrington, G. & Ang, D. L. Linear dichroism as a probe of molecular structure and interactions. *The Analyst* **141**, 6490–6498 (2016).
49. Tridgett, M. *et al.* Linear dichroism of visible-region chromophores using M13 bacteriophage as an alignment scaffold. *RSC Advances* **8**, 29535–29543 (2018).
50. Wirtz, A. C., Hofmann, C. & Groenen, E. J. J. Stretched Polyethylene Films Probed by Single Molecules. *ChemPhysChem* **12**, 1519–1528 (2011).
51. Razmkhah, K., Chmel, N. P., Gibson, M. I. & Rodger, A. Oxidized polyethylene films for orienting polar molecules for linear dichroism spectroscopy. *Analyst* **139**, 1372–1382 (2014).
52. Stopar, D., Spruijt, R. B., Wolfs, C. J. A. M. & Hemminga, M. A. Protein–lipid interactions of bacteriophage M13 major coat protein. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1611**, 5–15 (2003).
53. What is Fluorescence Spectroscopy? https://www.horiba.com/en_en/technology/spectroscopy/fluorescence-spectroscopy/what-is-fluorescence-spectroscopy/.
54. Nagashima, K. & Velan, S. S. Understanding the singlet and triplet states in magnetic resonance: Singlet and Triplet States in Magnetic Resonance. *Concepts Magn. Reson.* **42**, 165–181 (2013).
55. Fluorescence Spectroscopy (The Basics) - JASCO. *JASCO Inc.* https://jascoinc.com/learning-center/theory/spectroscopy/fluorescence-spectroscopy/.
56. Granite. Jablonski Diagram | What is it? *Edinburgh Instruments* https://www.edinst.com/blog/jablonski-diagram/.
57. Valeur, B. Molecular Fluorescence. in *digital Encyclopedia of Applied Physics* (ed. Wiley-VCH Verlag GmbH & Co. KGaA) 477–531 (Wiley-VCH Verlag GmbH & Co. KGaA, 2009). doi:10.1002/3527600434.eap684.
58. Lakowicz, J. R. *Principles of Fluorescence Spectroscopy*. (Springer, 2006).

59. Albani, J. R. *Structure and dynamics of macromolecules: absorption and fluorescence studies*. (Elsevier, 2004).

60. Dragan, A. I. *et al.* SYBR Green I: Fluorescence Properties and Interaction with DNA. *J Fluoresc* **22**, 1189–1199 (2012).

61. SYBR® Green Based Quantitative PCR. *Sigma-Aldrich* https://www.sigmaaldrich.com/life-science/molecular-biology/pcr/quantitative-pcr/sybr-green-based-qpcr.html.

62. SYBR Green for qPCR | Bio-Rad. https://www.bio-rad.com/featured/en/sybr-green-for-qpcr.html.

63. Zipper, H., Brunner, H., Bernhagen, J. & Vitzthum, F. Investigations on DNA intercalation and surface binding by SYBR Green I, its structure determination and methodological implications. *Nucleic Acids Res* **32**, e103 (2004).

64. Carter, J. G. *et al.* Ultrarapid detection of SARS-CoV-2 RNA using a reverse transcription–free exponential amplification reaction, RTF-EXPAR. *Proc Natl Acad Sci USA* **118**, e2100347118 (2021).

65. Coskun, O. Separation Techniques: CHROMATOGRAPHY. *North Clin Istanbul* (2016) doi:10.14744/nci.2016.32757.

66. Instruments, S. S. High Performance Liquid Chromatography (HPLC) Basics. https://www.ssi.shimadzu.com/ (2020).

67. Snyder, L. R., Kirkland, J. J. & Glajch, J. L. *Practical HPLC Method Development*. (2012).

68. Huber, U. & Majors, R. E. Principles in preparative HPLC. 85.

69. Ortega, F., Velez, E. & Somanathan, R. Synthesis and Use of Reverse-Phase Silica Gel for HPLC in Undergraduate Chemistry. *J. Chem. Educ.* **73**, A26 (1996).

70. Cramer, H., Finn, K. & Girindus, E. Purity Analysis and Impurities Determination by Reversed-Phase High- Performance Liquid Chromatography. in *Handbook of Analysis of Oligonucleotides and Related Products* (ed. Srivatsa, G.) 1–46 (CRC Press, 2011). doi:10.1201/b10714-2.

71. Molnár, I. & Horváth, C. Reverse-phase chromatography of polar biological substances: separation of catechol compounds by high-performance liquid chromatography. *Clinical Chemistry* **22**, 1497–1502 (1976).

72. Gross, J. H. *Mass Spectrometry: A Textbook*. (Springer International Publishing : Imprint: Springer, 2017). doi:10.1007/978-3-319-54398-7.

73. Hoffmann, E. de & Stroobant, V. *Mass spectrometry: principles and applications*. (J. Wiley, 2007).

Chapter 3: Detection of isothermally amplified DNA using linear dichroism spectroscopy

3.1 Introduction

The detection of deoxyribonucleic acid (DNA) is a method for the identification of infection-causing pathogens. Due to the small concentration of genetic material in pathogens compared to host genetic material, most nucleic acid tests require an amplification step to increase the quantity of DNA to a more detectable level. Many of these tests employ the polymerase chain reaction (PCR), however, recently the emergence of isothermal amplification techniques has challenged the dominance of PCR due to its rapid nature (<1 hour).

This chapter will discuss and compare some of the current methods used for rapid amplification of DNA. The latter part of the chapter will then describe the results from the optimisation of our chosen isothermal amplification technique (EXPAR), and the subsequent detection of DNA *via* linear dichroism (LD) spectroscopy using DNA-tagged M13 Bacteriophage.

3.2 Isothermal Amplification

The most adopted method for the amplification of DNA is PCR; however, due to the requirement of thermocycling and a lengthy amplification process, this makes the technique unsuitable for point of care detection. Isothermal amplification on the other hand offers a method that substitutes thermocycling with a constant temperature. One of the most widely adopted method is loop mediated isothermal amplification (LAMP) (Chapter 1.4.3.4.2) which does reduce the overall time of the reaction albeit with increased complexity of the system. There are other isothermal methods which could similarly provide point of care amplification without the complexity required for LAMP.

3.2.1 Strand Displacement Amplification

Strand displacement amplification (SDA) is a method which uses a DNA polymerase and a restriction enzyme to amplify DNA strands exponentially.[1] All amplification methods employ the use of deoxynucleoside triphosphates (dNTPs) as building blocks for the DNA sequence. In the case of SDA, deoxycytidine triphosphate (dCTP) is substituted with deoxycytidine thiotriphosphate (dCTPαS).[1] The restriction enzyme is unable to cut bases linked *via* a phosphorothioate groups, effectively turning the restriction enzyme into a nicking enzyme, as it is only able to cut an unmodified strand.[2] SDA itself is performed using four primers, B1, B2, S1 and S2.[1–4] S1 and S2 are composed of a sequence containing a restriction enzyme recognition site at the 5´ end, and a target recognition site at the 3´ end; B1 and B2 on the other hand contain no enzyme recognition sites.[1–4] This technique enables for the amplification of dsDNA at 60 °C, with the reaction beginning when S1 and B1 hybridise to the target. The DNA polymerase extends S1 and B1, and as B1 is extended the S1 extension product (S1-ext) is displaced (**Figure 3.1**).

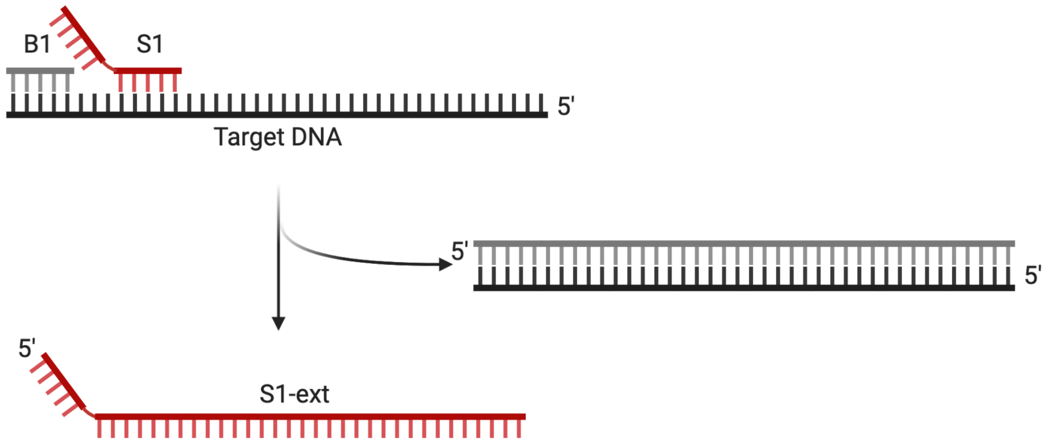


Figure 3.1: *Primer S1 and B1 anneal to the denatured target DNA. Extension of B1 displaces the product of extension S1-ext.*

S1-ext possesses complementary regions for the primers S2 and B2. The simultaneous extension of both primers produces the displaced extension product S2-ext, to which another S1 primer can bind and extend (**Figure 3.2**).

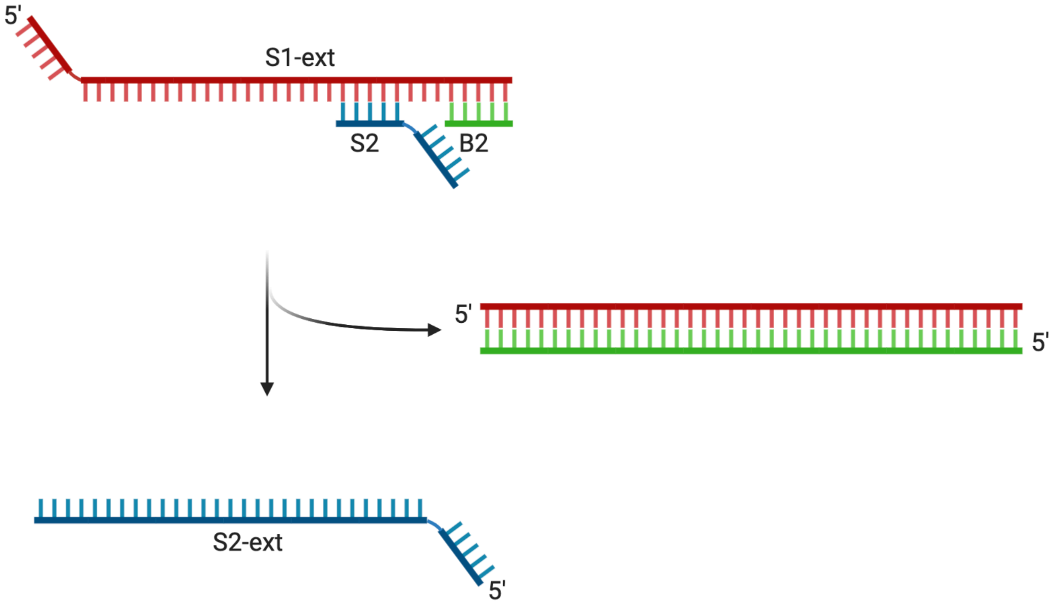


Figure 3.2: *S1-ext acts as a template for primers B2 and S2. Extension of B2 displaces product of extension S2-ext.*

The extension of S1 annealed to S2-ext produces a duplex containing a hemiphosphorothioate enzyme recognition site. The S1 strand is nicked, leaving the S2-ext strand intact. DNA polymerase extends the nicked S1 strand, displacing the longer oligonucleotide sequence.[1–4] A S2 primer then binds to the displaced strand and is extended, thus allowing the SDA cycle to being (**Figure 3.3**).

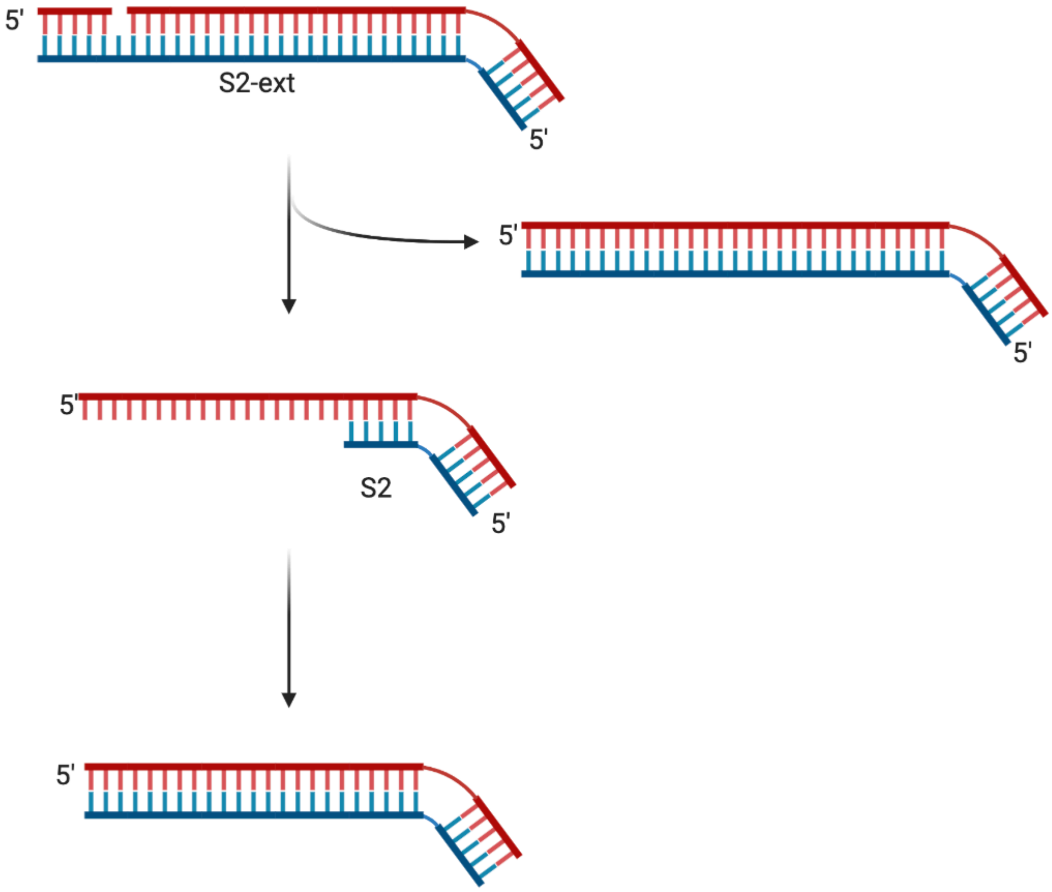


Figure 3.3*: The duplex formation generates a hemiphosphorothioate restriction site, the restriction enzyme nicks the unmodified strand of the duplex. Following which the short 3´ sequence is able to amplify and displace the longer 5´ end. The displaced sequence acts as a template for S2. Duplex formation occurs creating the template for the SDA cycle*

The SDA cycle begins when the S2-ext is nicked at the hemiphosphorothioate enzyme recognition site, releasing a displaced target strand (T2). During each round of the SDA cycle, the S1 binds to T2 leaving an overhang (**Figure 3.4** i). DNA polymerase extends both sequences to create a full duplex containing a hemiphosphorothioate recognition site (**Figure 3.4** ii), that is subsequently nicked (**Figure 3.4** iii), and displaces following extension (**Figure 3.4** iv). The displaced product the complementary of T2 (T1). As a result of this, T1 binds with a new S2 binder creating more T2 continuously (**Figure 3.4** i).[1–4]

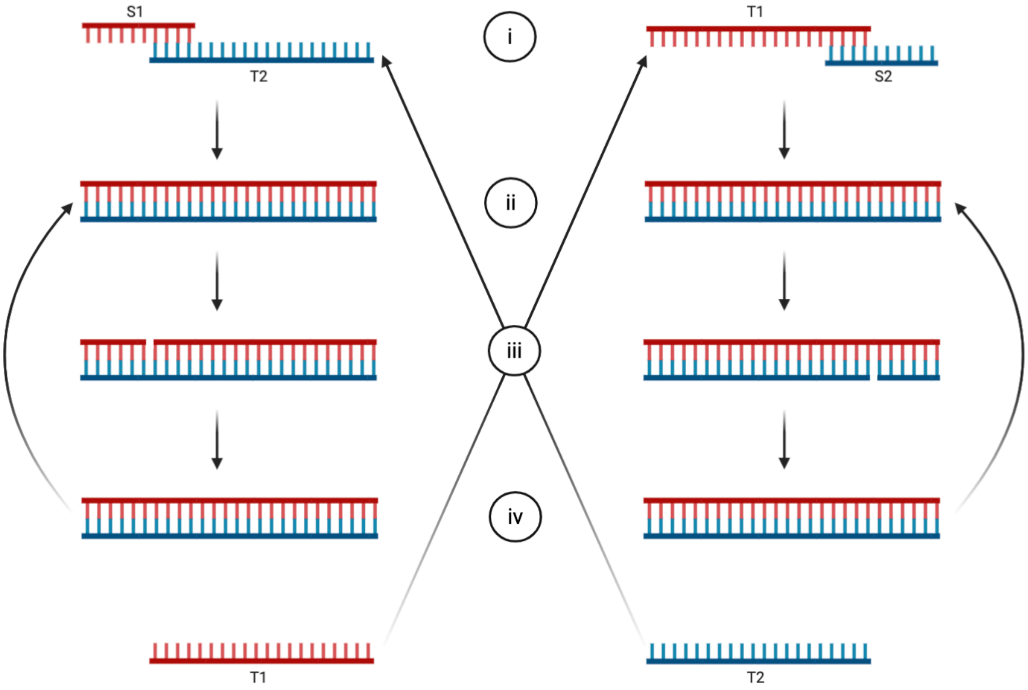


Figure 3.4: *The SDA cycle. The final stage of the previous steps (**Figure 3.3**) starts at stage (ii). (i) Target strand (T1 and T2) acts as a template for primers (S2 and S1 respectively). The primers 3´ end binds to the targets 3´ end leaving the 5´ end of the primer overhanging. (ii) DNA polymerase extends both the primer and the target to create a full duplex with a hemiphosphorothioate binding site. (iii) The restriction enzyme nicks the unmodified sequence. (iv) Extension of the short primer following nicking displaces a strand that can go on to act as a new trigger sequence. Adapted from Walker et al.*[1]

SDA is a highly sensitive and specific amplification technique that can be detected using molecular beacons.[5] However, the nature of technique means only shorter sequences are able to be amplified (50 – 120 bp) in a longer time than PCR. Furthermore, SDA is a complicated technique to design and optimise.[6,7]

3.2.2 Helicase-Dependent Amplification

Helicase-dependent amplification (HDA) works through the employment of two enzymes, a DNA polymerase and a DNA helicase.[8,9] This process beings when a duplex DNA is unwound and separated by DNA helicase, in the presence of adenosine triphosphate (ATP). Following separation, single stranded binding protein (SSB) binds to the single DNA strands.[8–11] Specific primers are able to anneal to the 3´ ends of the ssDNA sequence. DNA polymerase then extends the primers to form two new DNA duplexes identical to the original. The newly generated DNA sequences are then subjected to the same process, thus causing an exponential amplification the target sequence (**Figure 3.5**).[8–11]

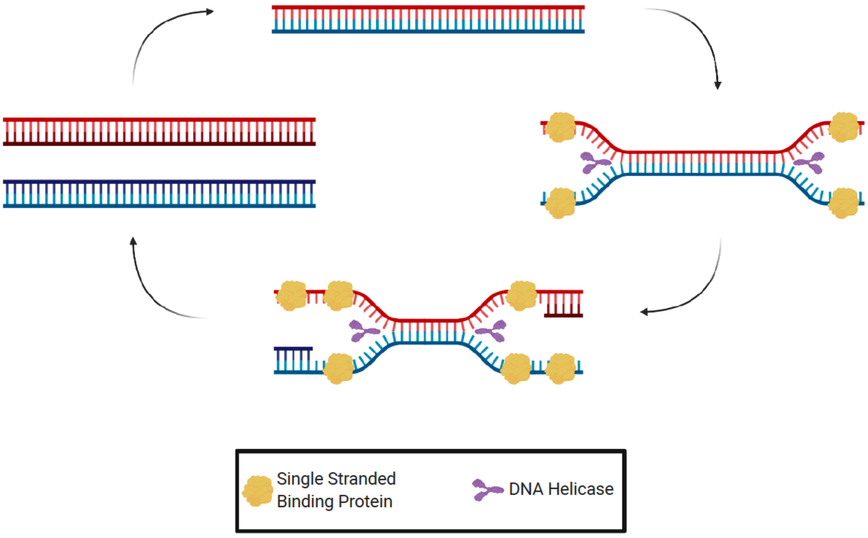


Figure 3.5*: The HDA cycle. dsDNA is unwound and separated by DNA helicase. SSB (Yellow) binds to the displaced ssDNA. Sequence specific primers bind to the displaced DNA. DNA polymerase extends the primers to produce two new sequence of dsDNA which are able to be used to begin the cycle anew.*

HDA offers the advantage of a simpler design than some other amplification methods. Furthermore, the ability to visualise amplification using fluorescence enables real-time monitoring of reaction progress. However, HDA is limited by the possibility of amplification inhibition or non-specific amplification as a result of primer mispairing.[12,13]

3.2.3 Rolling Circle Amplification

Rolling circle amplification (RCA) uses a DNA polymerase to continually amplify a circular sequence of DNA.[14] A circular sequence, such as a plasmid, is the ideal template; however through the use of a padlock probe a linear sequence of DNA can be circularised. A padlock probe is a sequence that anneals to both the 3´ and 5´ end of an oligonucleotide, thus creating a link between each end, and, as a result a circular sequence (**Figure 3.6**).[15]

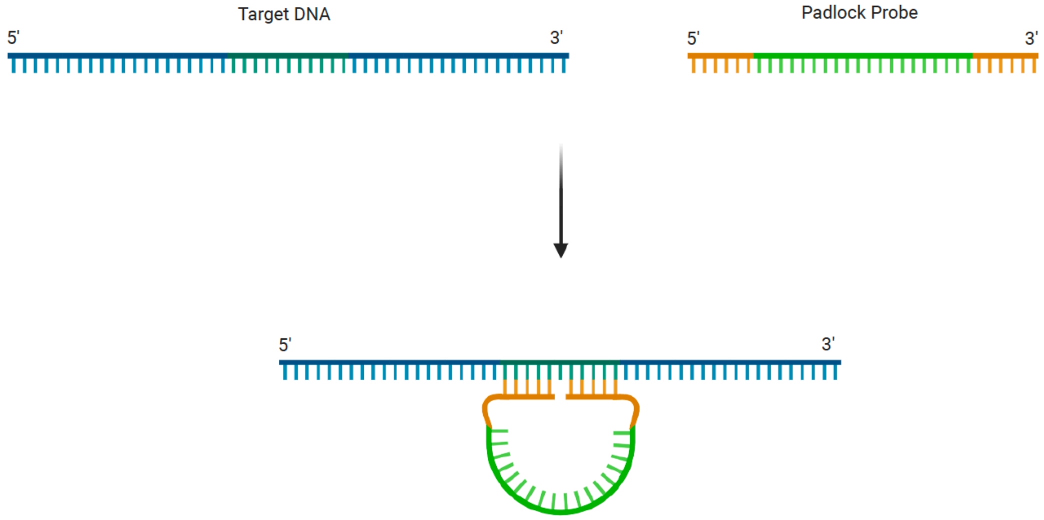


Figure 3.6: *Padlock probe formation. The terminal regions (Yellow) of padlock DNA probe are complementary to the target region of the double stranded DNA (Green). The annealing of the probe to the target region creates a circular sequence of DNA.*[15]

The amplification process begins when a primer anneals to the circular DNA template (**Figure 3.7 A**). DNA polymerase extends the primer to create a DNA duplex (**Figure 3.7 B**) and upon reaching the priming site, the DNA polymerase displaces the original strand to continue amplification (**Figure 3.7 C**).[14,16–18] Post amplification, DNA can be detected *via* fluorescence although it has also been reported that the incorporation of gold nanoparticles and quantum dots have been used for visualisation.[19] A major advantage of RCA over other methods is the ability to hyper branch. Through the inclusion of multiple primers for different areas of the circular DNA, it is possible to produce multiple sequences simultaneously thus creating more DNA in a shorter period of time. A further primer can be added that is complementary to the product sequence, thus leading to amplification to occur on the newly synthesised strand creating hyperbranched DNA (**Figure 3.7 E**).[20]

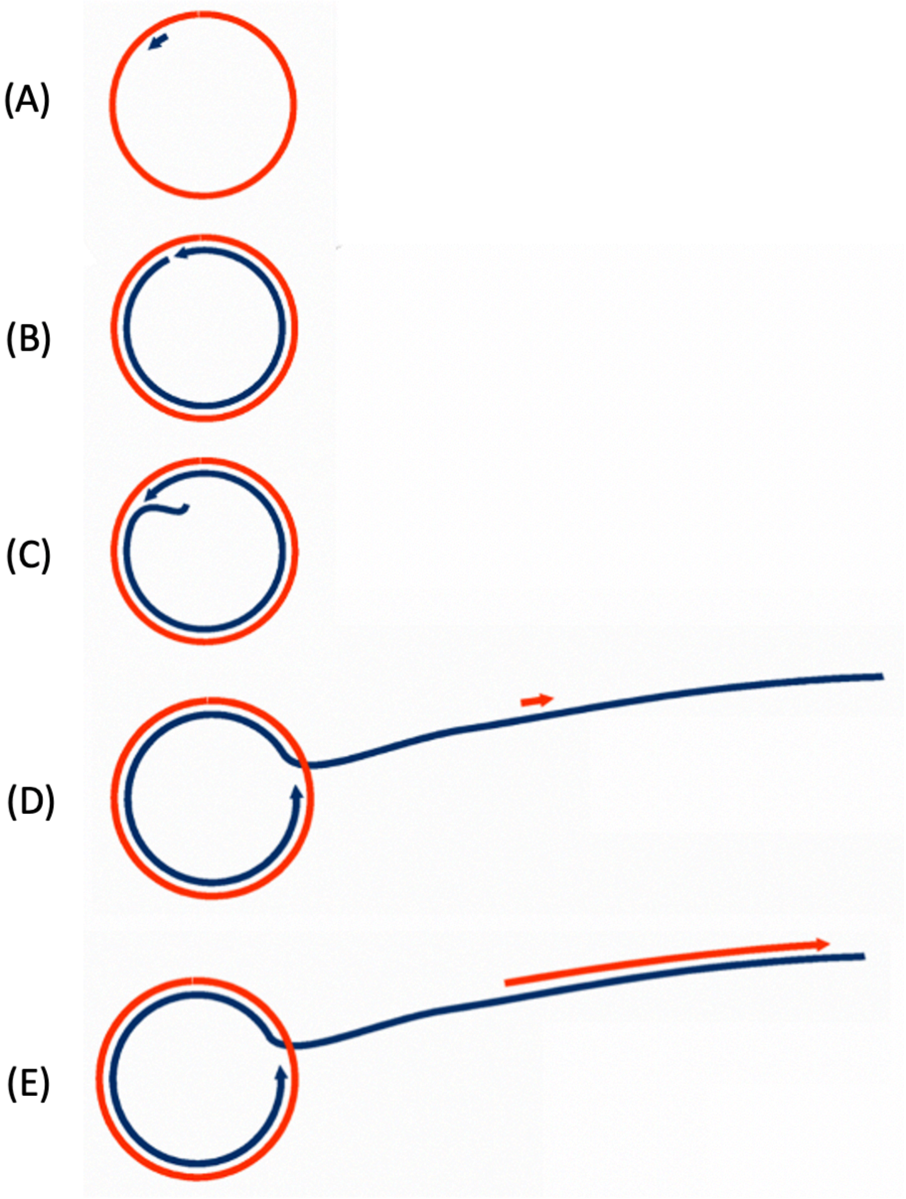


Figure 3.7: *Rolling Circle Amplification. (A) The primer (Blue) anneals to the circular target (Red). (B) Amplification of the primer creates a circular duplex. (C) Upon reaching the priming site, the DNA polymerase displaces the original strand to continue amplification. (D) The displaced strand can be primed by a second primer. (E) The second primer is amplified. Adapted from Jeske et al.* [21]

RCA is not without its limitations - primer design is vital and can pose a challenge, as the selection of inadequate binding sites reduces the effectiveness of enzymes. Furthermore, long storage of RCA products leads to non-specific crosslinking, hindering the ability for mass production of DNA.[22]

3.2.4 Exponential Amplification Reaction

The exponential amplification reaction (EXPAR) uses a template sequence, a DNA polymerase, and a nicking enzyme. EXPAR uses the target DNA sequence as the short primer, this is designated as trigger X. The trigger is amplified using an exponential template (X´-X´) consisting of two repeating sequences complementary to X, separated by the recognition site for a nicking enzyme (**Figure 3.8**)[23–26] which only cuts the top strand of DNA, leaving the template untouched (**Figure 3.9**).[27]

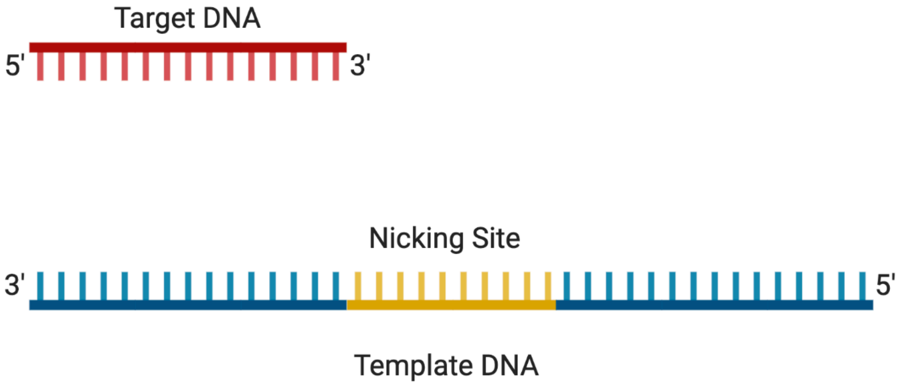


Figure 3.8: *(Top) The Trigger sequence for an EXPAR reaction that will be amplified. (Bottom) The exponential template, composed of two X´ sequences, complementary to the trigger, separated by a nicking enzyme recognition site*

The amplification itself begins with the hybridisation of the trigger with the template, following which a DNA polymerase extends the trigger DNA forming a full duplex. The now double stranded recognition site is nicked by the enzyme, generating a new trigger sequence.

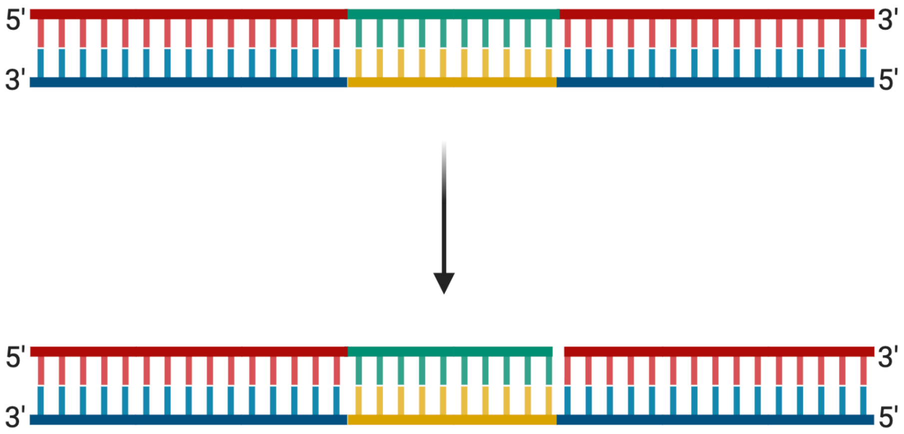


Figure 3.9: *Upon hybridisation of Trigger to template, the sequence is extended with a DNA polymerase, completing the nicking site, which is consequently nicked allowing the synthesised X to be released*

Consequently, the cut product DNA is displaced from the template DNA, allowing the DNA polymerase to elongate the trigger DNA again. As the released DNA is identical to the original trigger, this can go on to prime another template, creating yet more product DNA, and hence exponentially producing more trigger DNA (**Figure 3.10**).[23,24,28,29]

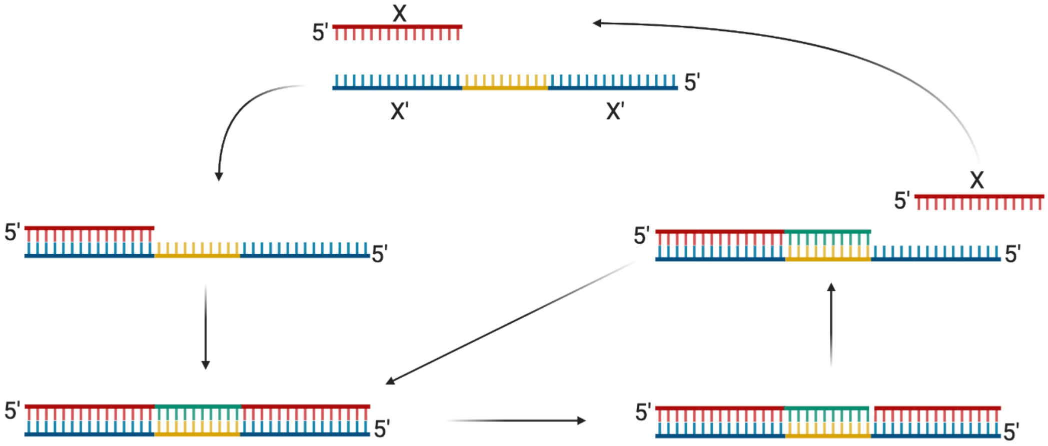


Figure 3.10*: The exponential EXPAR cycle, starting with the annealing of Tripper to template, followed by extension, nicking and finally release. The released DNA is able to go on and prime further templates. Adapted from Nie et al.*[29]

The inclusion of a linear template (X′–Y ′), composed of a sequence complementary to the trigger and a secondary sequence separated by an enzyme nicking recognition site, enables the generation of the oligonucleotide Y, that if required, can be used as a universal recognition sequence for DNA binding detection system (**Figure 3.11**).[24,27,28]

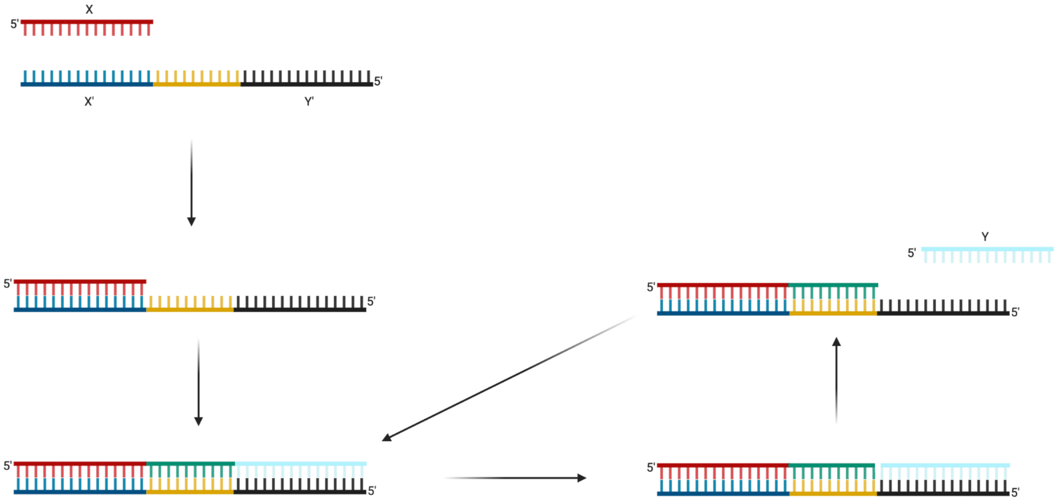


Figure 3.11: *The linear EXPAR cycle, starting with the annealing of Tripper to template, followed by extension, nicking and finally release. The released DNA is unable to prime DNA but can be detected using another system. Adapted from Nie et al.*[29,30]

EXPAR is an extremely rapid amplification process, producing $10^6 - 10^8$ copies in less than 20 minutes.[31] It is however a difficult amplification technique to master, as the reaction temperature must be a balance between each enzymes optimal range. Additionally, EXPAR can produce non-specific amplification, this is a result of an unknown reaction, however it is believed to be causes by a template interacting with another template.[32]

3.2.5 Summary of Nucleic Acid Amplification

The current methods discussed above offer a variety of approaches for the amplification of DNA, with the key points summarised in **Table 3.1**. PCR is the most adopted method of amplification; however, the requirement of a thermocycler (i.e., repeated heating and cooling steps) makes the system unfavourable for point of care (POC) testing. While all the isothermal techniques discussed successfully amplify DNA, differences become apparent when the reactions conditions are observed. Most of them require multiple enzymes in order to function and therefore a similar temperature window for optimum activity. Due to the intended application of this work (POC testing), a shorter reaction time is the most desirable, which makes RCA and SDA unfavourable as their reaction times take at least an hour.

LAMP is currently favoured amongst multiple research groups due to the good specificity and sensitivity of the assay, however, primer design is extremely complicated. HDA and EXPAR are extremely promising, being able to amplify short sequences of DNA, however, EXPAR has the advantage as it is able to produce DNA in less than half an hour. Additionally, the product of EXPAR is ssDNA whereas HDA is dsDNA, the amplification of ssDNA, makes an ideal target for our assay.

Table 3.1: *A summary of the key features of nucleic acid amplification techniques. Table adapted from Zhao et al.*[33]

	PCR	**LAMP**	**EXPAR**	**SDA**	**HDA**	**RCA**
Target	DNA & RNA	DNA & RNA	Short DNA	DNA	DNA	DNA & RNA
Product	DNA	DNA	DNA	DNA	DNA	DNA
Number of Primers	2	4	1	4	2	2
Number of Enzymes	1	1	2	2	2	1
Enzyme Class	DNA Polymerase	DNA Polymerase	DNA Polymerase, Nicking Endonuclease	DNA Polymerase, Restriction Endonuclease	DNA Polymerase, DNA Helicase	DNA Polymerase
Reaction Temperature (°C)	60 – 95	~ 65	~ 50	37	~ 37	~ 60
Reaction Time (h)	1.5	< 1	< 0.5	2	0.5 – 2	1 – 4

3.3 Project Aims

Previous work in the Tucker and Dafforn groups demonstrated the ability to detect genetic material of microorganisms using a M13 Bacteriophage – LD system.[34] However, the limit of detection was still too high for a rapid point of care (POC) test. This project was a proof of concept, to determine if it was possible to combine an LD assay system with a method to rapidly amplify DNA targets. The method chosen to amplify DNA was EXPAR, a technique capable of amplifying short sequences of DNA at a fast rate. The aims to achieve this were:

1) Design and optimise an EXPAR reaction to produce a detectable quantity of DNA before unspecific amplification occurs,
2) Determine that the oligonucleotide being amplified were those of the desired sequence.
3) Design a M13 Bacteriophage probe that could be applied for detecting the output of an EXPAR *via* LD.

3.4 Results and Discussion

3.4.1 Sequence design

This chapter focuses on the combination of an amplification and detection system for the identification of DNA. For this study we opted for the detection of the bacterial infection *Chlamydia Trachomatis*, one of the most common sexually transmitted infections, afflicting over 61 million people per year.[35] We envisioned that our system could amplify a short conserved sequence of a *Chlamydia* plasmid (**Trigger X**) using a two-stage EXPAR system. The second stage would generate sequence (**Reporter Y**) which could be detected using the M13 Bacteriophage–LD system (**Figure 3.12**).

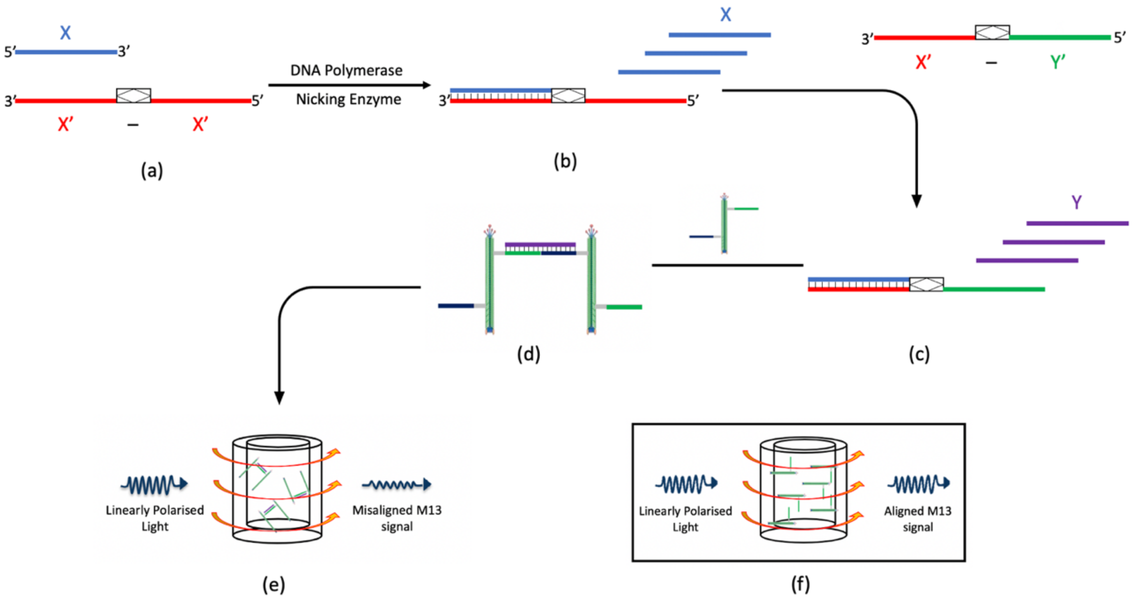


Figure 3.12: *Diagrammatic representation of oligonucleotide amplification using EXPAR and detection using linear dichroism spectroscopy. (a, b) Exponential amplification of* ***Trigger X*** *in the first stage of the EXPAR reaction.* ***Trigger X*** *primes the X´-X´ template. DNA polymerase and the nicking enzyme lead to the production of more* ***Trigger X*** *through cycles of amplification and nicking. (c) The conversion of* ***Trigger X*** *into* ***Reporter Y****. The X´-Y´ template enables* ***Trigger X*** *to prime and produce* ***Reporter Y****. (d) Detection of* ***Reporter Y****. The addition of DNA-conjugated M13 Bacteriophage causes hybridization when in the presence of* ***Reporter Y****. (e) When hybridized, the M13 Bacteriophage cannot align under sheer flow due to aggregation, which results in a decreased LD signal, indicating the presence of target DNA. (f) When no* ***Reporter Y*** *is present to hybridize to the DNA-conjugated M13 Bacteriophage, a large LD signal remains under sheer flow full alignment.*

Reporter Y could be any sequence, so long as it would not interact with the amplification reaction itself. Therefore the reporter sequence selected was the same as the one Tan *et al* had previously demonstrated to work effectively.[23,28] **Trigger X** on the other hand is a more complicated design process. Initially, the trigger must be found within the *Chlamydia Trachomatis* genome but not found within any other species. Secondly, the sequences must possess a T_m value that is close to the reaction temperature. Finally, work by Qian *et al* has demonstrated that the DNA base sequence at both ends of the trigger plays a key part to the success of the EXPAR reaction. The base sequence 5´-AGGGG–CCAC-3´ was found to improve EXPAR, with the bases in-between G and C having less of an impact. We have therefore used the findings from this work to design our trigger sequence.[23] Upon examination of the CDS2 gene of the pSotonF3 plasmid, 4 sequence which best matched the above constraints were selected (**Table 3.2**). Analysis of the 4 sequences showed a differences between the T_m values of the sequences. Sequence 3 shows a T_m of 55.4 °C, as the EXPAR reaction was to proceed at 50 °C, the Trigger T_m was deemed too high for EXPAR to proceed as intended and therefore this sequence was discounted. Meanwhile Sequence 4 shows a T_m of 36.0 °C, this was determined to be too low for the reaction to proceed and was also discounted.

Table 3.2: *Potential Trigger sequences identified within the Chlamydia Trachomatis cryptic plasmid and their respective melting temperatures.*

Sequence	Trigger (5´-3´)	T_m
1	CCG GGA TTG GTT GAT	48.2 °C
2	GGG TGC TCA GAC TCC	51.8 °C
3	GCA AAT CGC CCG CAC	55.4 °C
4	AGG GAT TTT ATC TTT	36.0 °C

From the remaining sequences, sequence 1 was designated as the target of choice (***Chlamydia* Trigger**) and sequence 2 was designated as a control sequence (**Control Trigger**), which would not be complementary to any template and therefore should not be amplified (**Table 3.3**).[36,37] Following the selection of both trigger and reporter we were able to design the two template sequences. An exponential template (***Chlamydia*´- *Chlamydia*´**), involved in producing more trigger to prime templates, was composed of two regions complementary to ***Chlamydia* Trigger**, separated by a nicking enzyme recognition site. The linear template (***Chlamydia*´-Reporter´**), involved in producing the reporter sequence, was composed of a region complementary to ***Chlamydia* Trigger** and a region complementary to **Reporter** separated by a nicking enzyme recognition site. During the design process, each template was tailored to feature an additional guanine (G) base at the 5´ end of the enzyme recognition site, this produces an additional cytosine (C) base at the 5´ end of newly synthesised ***Chlamydia* Trigger** and **Reporter**. While this has no adverse effect on the binding of the trigger to the template, it does enable the identification and differentiation of oligomer product to oligomer trigger in mass spectrometry. The DNA strands for conjugation to M13 Bacteriophage were each designed to be fully complementary to the synthesised **Reporter** when they are combined, with a thiol group at either their 5´ or 3’ ends. (**Table 3.3**) All sequences used in this chapter were purchased from Sigma Aldrich.

Table 3.3: *DNA sequences used throughout this study. Grey highlighted, Nt.BstNBI recognition site, red letters; complementary sequence to* ***Chlamydia Trigger****, blue letters;* ***Chlamydia Trigger****, Green letters; complementary sequence to* ***Reporter****, Purple letters;* ***Reporter****, yellow highlighted region; additional G base. Thiolated DNA was purchased in the disulfide form and reduced prior to conjugation.*

Name	Sequence (5´ – 3´)	Extinction Coefficient (L/mol·cm)
Probe 1 Y´	HS – CTG GCG CT	66,200
Probe 2 Y´	TGA TGG TAG – SH	84,500
Control Trigger	GGG TGC TCA GAC TCC	137,300
***Chlamydia* Trigger**	CCG GGA TTG GTT GAT	143,000
Template *Chlamydia´*- *Chlamydia´*	ATC AAC CAA TCC CGG GTG AGA CTC TAT CAA CCA ATC CCG G	383,900
Template *Chlamydia´*-Reporter´	CTG GCG CTT GAT GGT AGT GAG ACT CTA TCA ACC AAT CCC GG	387,800
Reporter	TAC CAT CAA GCG CCA G	153,900

3.4.2 M13 Bacteriophage Propagation, bioconjugation and characterisation

Following the acquisition of the DNA probes, M13 Bacteriophage was required in order to produce probes for LD sensing. M13 Bacteriophage was propagated through the infection of *Escherichia coli (E. coli)* with a small quantity of the virus. The M13 Bacteriophage hijacks the cellular functionality of the bacterium, to produce numerous progeny M13 Bacteriophage without triggering cell death of the *E. coli* host. Purification and concentration of M13 Bacteriophage from the bacterial broth is carried out to using a well-defined method of extraction.[38] Following collection of M13 Bacteriophage, its concentration was determined using Ultraviolet Visible (UV-Vis) spectroscopy (**Figure 3.13**) (Chapter 6.3.5.1).

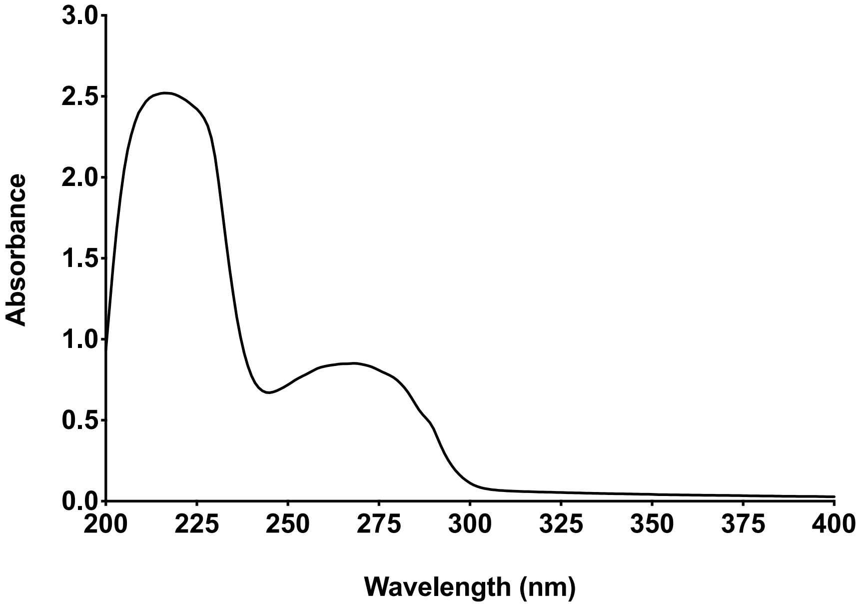


Figure 3.13: *A characteristic UV-Vis spectrum of M13 bacteriophage. M13 Bacteriophage was characterised in 50 mM potassium phosphate buffer pH 8.0 at room temperature.*

Following propagation, M13 Bacteriophage probes were produced through the conjugation of probes 1 and 2 (**Table 3.3**) to a single M13 Bacteriophage molecule using a method developed in the Tucker and Dafforn groups.[38,39] Firstly, the cross linker SMCC was used to form a stable amide bond between the amine group on pVIII protein and the *NHS*-ester of SMCC. The formation of this bond leaves the maleimide group available to react with the thiol tagged DNA, which itself is formed by reduction of the S-S bond in **Probe 1** and **Probe 2** (**Figure 3.14**).[39,40]

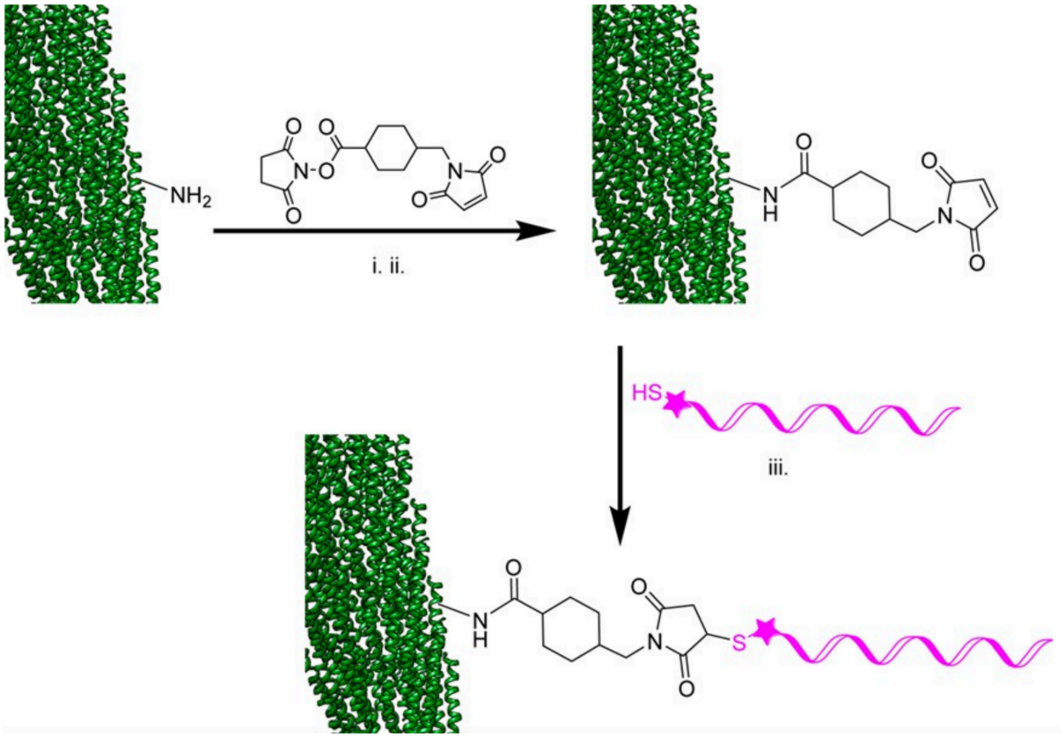


Figure 3.14: *A reaction scheme showing how SMCC can be used as a crosslinker, the initial reaction between the amine of the M13 Bacteriophage and the NHS ester of the SMCC molecule forms an amide bond. A thiol-linked DNA reacts with the C-C double bond in the maleimide which results in the formation of the conjugation between the protein and the peptide. Image taken from Carr-Smith et al.*[39]

Following bioconjugation, an LD spectrum was measured in order to determine that the structure of the M13 Bacteriophage remained intact (**Figure 3.15**).

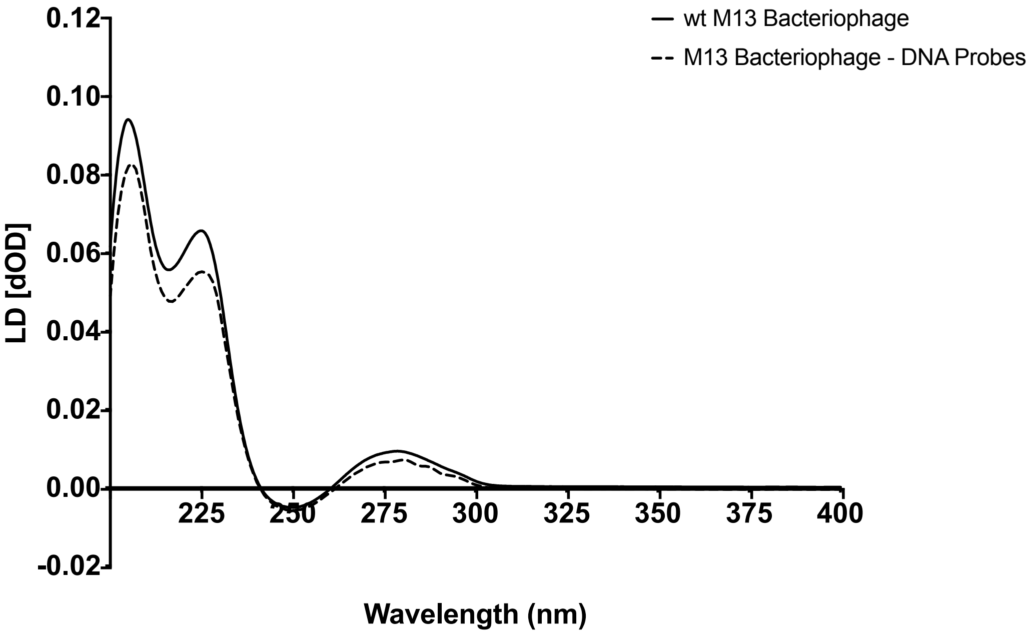


Figure 3.15: *An LD spectrum of wtM13 Bacteriophage (Solid Line) and an LD spectrum of M13 Bacteriophage after conjugation to DNA probes (Dashed Line). M13 Bacteriophage was characterised in 1x Phosphate Buffered Saline (PBS) pH 7.4 at room temperature (n = 3).*

3.4.3 DNA Binding Studies

As shown in **Figure 3.16** the presence of target DNA causes multiple M13 Bacteriophage probes to bridge together in a sandwich-type arrangement, causing aggregation and thus a change in the hydrodynamic behaviour of the M13 Bacteriophage; in other words, aggregation removes its alignment in flow and as a result causes a decrease in the LD signal. These changes contrast with conventional methods of characterising target binding such as mass spectrometry and UV-Vis spectrometry, which would be ineffective due to the size of M13 Bacteriophage and the lack of any changes to the UV-Vis spectrum.

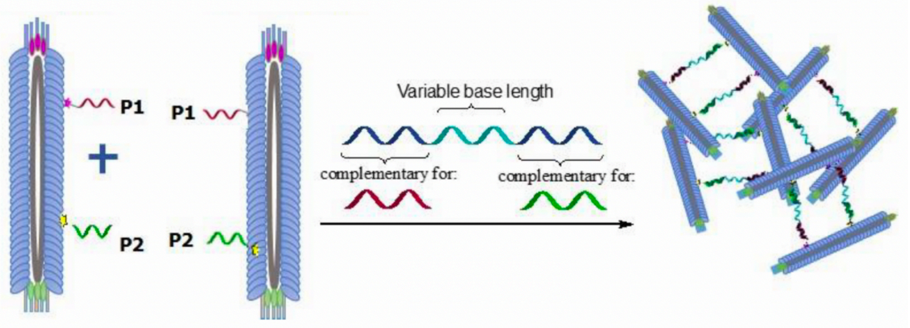


Figure 3.16: *The 2 probe, 1 phage sandwich assay. The left hand shows the M13-probes without target present; the right hand shows the M13-probes when the correct target is present. Image taken from Little.*[38]

Through the use of LD it is possible to record the alignment of M13 Bacteriophage probes in a micro Couette cell.[41] Using **Equation 3.1** the volume of M13 Bacteriophage probe required to obtain an LD signal of 0.01 dOD from a 100 µL sample can be calculated.

$$Volume\ of\ Probe = \frac{100\ \mu\text{L}}{\left(\frac{LD_{225} \times df}{0.01}\right)}$$

Equation 3.1: *The equation used in order to determine the quantity of M13 Bacteriophage – DNA probe required in order to achieve an LD value of 0.01 at 225 nm*

Using the above equation, the quantity of M13 Bacteriophage probe in a solution can be standardised to ensure that the LD signal of probe without target will always have an LD signal of 0.01. For sample analysis 10 µL of the buffer solution is be substituted for sample, this will have no effect on the overall concentration on probe in solution, therefore any change in LD signal will be a result of target binding. Using the M13 Bacteriophage with no target DNA as a control (3 measurements), the mean LD signal at 225 nm and the corresponding standard deviation could be calculated; from this we were able to determine that a significant change in alignment was present when the mean LD value at 225 nm of the probe was greater than 5 standard deviations (σ) below than the control mean (**Figure 3.17**). A series of LD spectra were recorded in the presence of different amounts of complementary target **Reporter**, as also shown in **Figure 3.17**. As expected, a decrease in LD signal was observed, with a significant difference being observed at concentrations of 60 nM and above, up to 100 nM. As expected, addition of DNA targets ***Chlamydia*** **Trigger** and **Control Trigger** (100 nM) that were non-complementary to M13-conjugated DNA strands **Probe 1** and **Probe 2** brought about no change in signal.

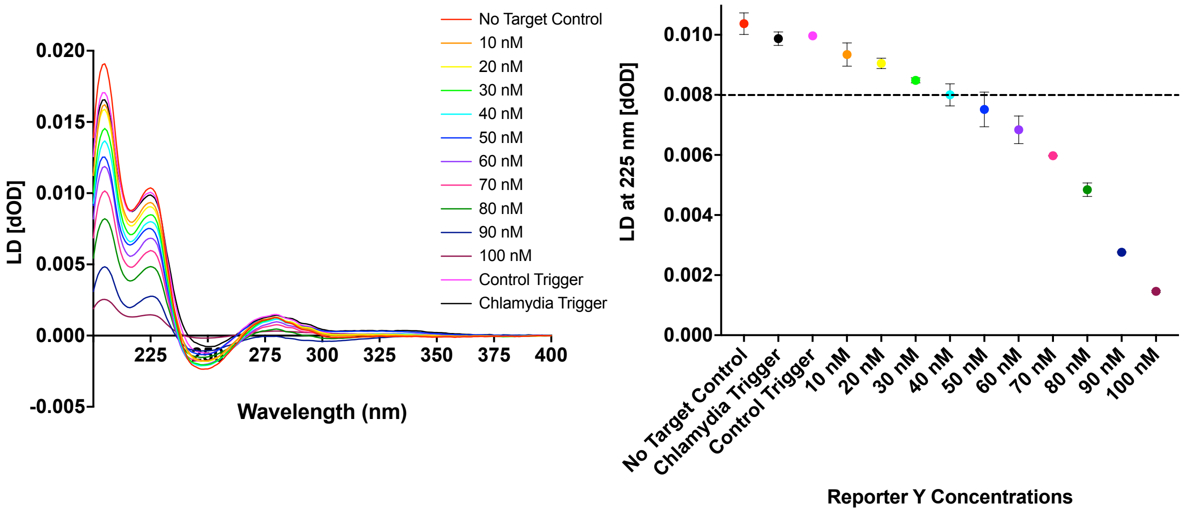


Figure 3.17: *(Left) Spectra produced from LD spectroscopy of M13 Bacteriophage – DNA probes with 0 to 100 nM of Reporter. (Right) LD values of M13 Bacteriophage – DNA probes with 0 to 100 nM of Reporter at 225 nm. Dashed line represents the 5σ threshold value. Concentration of Chlamydia Trigger and Control Trigger = 100 nM. Runs performed in triplicate (n = 3). Error bars in datasets are the standard deviations of the mean.*

3.4.4 Fluorescence Amplification Studies

With binding and sensing of the **Reporter** strand successfully demonstrated; the next stage was to optimise the amplification method that would produce the reporter strand. EXPAR is an extremely sensitive and rapid amplification technique,[24,26,42] however, this advantage of EXPAR is, unfortunately, also its drawback. The high sensitivity of EXPAR results in unspecific amplification occurring relatively quickly.[31,43] Therefore, the first task was to investigate this process using the intercalating dye, SYBR Green I, to monitor amplification *via* duplex formation. The intercalation between bases within dsDNA results in increased fluorescence emission compared to ssDNA,[44] a process commonly utilised as a read out for qPCR.[45] For EXPAR, rather than raising and then lowering the temperature, the temperature of the qPCR machine was kept constant at 50 °C. This allowed for the amplification of positive and negative samples to be monitored, differentiated, and ultimately extended to the largest possible degree.

As shown in **Table 3.3**, two DNA sequences were designed: ***Chlamydia* Trigger** and **Control Trigger**, the negative control. Using the method demonstrated in the seminal work by Van Ness *et al*[31] three sets of EXPAR reactions were performed, testing the effects of varying template concentration on EXPAR amplification:

1) ***Chlamydia* Trigger** (10 pM) against varying concentrations of **Template *Chlamydia´*– *Chlamydia´***
2) **C*ontrol* Trigger** (10 pM) against varying concentrations of **Template *Chlamydia´*– *Chlamydia´***
3) No trigger against varying concentrations of **Template *Chlamydia´* – *Chlamydia´***

Using ***Chlamydia* Trigger** against varying concentrations of **Template *Chlamydia´* – *Chlamydia´*** we would initially be able to monitor the ability of EXPAR to amplify small quantities of DNA. As expected, the results (**Figure 3.18**) displayed a large increase in signal for samples containing 100 nM of **Template *Chlamydia´* – *Chlamydia´***, with smaller signals appearing at later rise times for lower template concentrations. This is not an unexpected result as a larger template concentration would result in a greater likelihood of interaction with the trigger. As each amplification increases the quantity of ***Chlamydia* Trigger** in the reaction exponentially, we see the formation of a sigmoidal curve sooner. A template concentration of 25 nM results in a longer amplification rise time, however, the curve does appear to approach the exponential growth stage of the curve at 15 minutes, a significant improvement on the majority of other amplification methods. At less than 25 nM of **Template *Chlamydia´* – *Chlamydia´***, there is no visible amplification, however considering the small concentration of trigger this, again, is not unexpected.

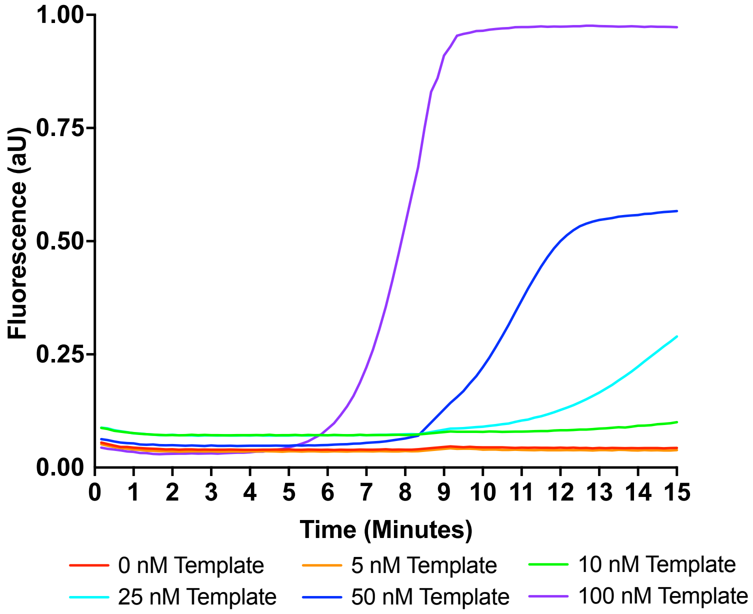


***Figure 3.18**: Relative normalised fluorescence data of SYBR Green I, obtained from EXPAR. The fluorescence produced from samples containing 10 pM of **Chlamydia Trigger**. Each sample possessed a different concentration of **Template Chlamydia´ – Chlamydia´**; 0 nM Template (Red) 5 nM template (Orange), 10 nM template (Green), 25 nM template (Turquois), 50 nM template (Blue), 100 nM template (Purple).*

These experiments clearly demonstrated that EXPAR was capable of amplifying 10 pM of DNA to detectable levels in a matter of minutes. The next study was to determine the effect of an alternative trigger; by substituting ***Chlamydia* Trigger**, for the non-complementary **Control Trigger** we are able to gauge the specificity of EXPAR. **Control Trigger** produces little to no amplification in the presence of **Template *Chlamydia´ – Chlamydia´*** (**Figure 3.19**).

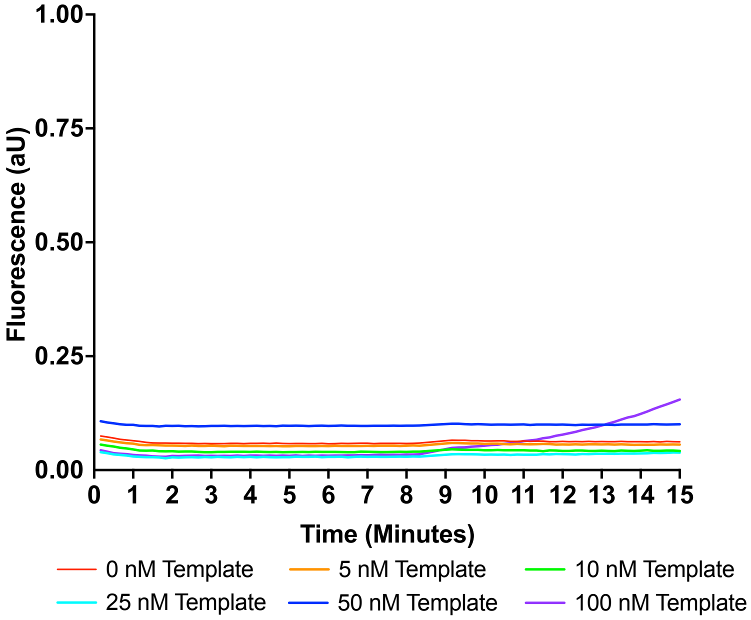


***Figure 3.19**: Relative normalised fluorescence data of SYBR Green I, obtained from EXPAR. The fluorescence produced from samples containing 10 pM of **Control Trigger**. Each sample possessed a different concentration of **Template Chlamydia´ – Chlamydia´**; 0 nM Template (Red) 5 nM template (Orange), 10 nM template (Green), 25 nM template (Turquois), 50 nM template (Blue), 100 nM template (Purple).*

Our final fluorescent study aims to assess the non-specific reaction that occurs in the absence of ***Chlamydia* Trigger**. We have previously seen (**Figure 3.19**) a small increase in duplex formation when a high concentration of **Template *Chlamydia´ – Chlamydia´*** is employed, yet this was still in the presence of a trigger, albeit a non-specific one. Performing the same studies as previously discussed yet with no trigger we were able to see (**Figure 3.20**) no amplification occurring in any sample except for 100 nM **Template *Chlamydia´ – Chlamydia´***. The results of these studies indicate that, while 100 nM produces some background amplification, this does not occur before 10 minutes compared to the 5 minutes amplification time of the fully complementary trigger . As none of the other samples displayed non-specific amplification within 15 minutes, we therefore opted to use 50 nM **Template *Chlamydia´ – Chlamydia´*** for the remainder of this work. This concentration produces detectable levels of amplification after 8 minutes and reduced the risk of detecting non-specific amplification.

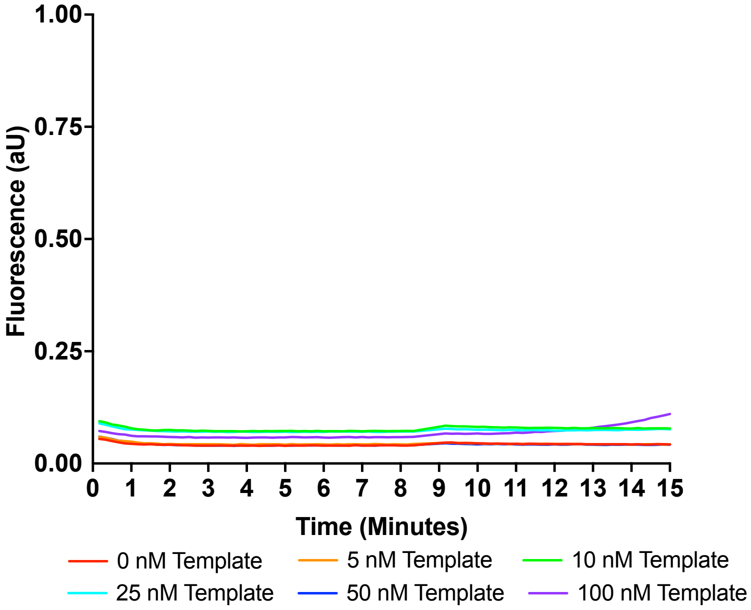


***Figure 3.20**: Relative normalised fluorescence data of SYBR Green I, obtained from EXPAR. The fluorescence produced from samples containing no trigger. Each sample possessed a different concentration of **Template Chlamydia´ – Chlamydia´**; 0 nM Template (Red) 5 nM template (Orange), 10 nM template (Green), 25 nM template (Turquois), 50 nM template (Blue), 100 nM template (Purple).*

While we were able to see amplification of a 10 pM trigger in under 10 minutes, we wanted to decrease the time while not affecting the reaction yield. The addition of additives in isothermal reactions has been shown to increase sensitivity, however they usually come at the cost of yield.[24] SSB has previously been employed in isothermal amplificant reactions as a method to prevent ssDNA sequences interacting and causing non-specific amplification.[11] Mok *et al* discovered that the inclusion of SSB can increase the specificity reaction while displaying no effect on the quantity of ssDNA produced. Guided by this the effects of various SSB concentrations were investigated (**Figure 3.21**). Using 50 nM of **Template *Chlamydia´* – *Chlamydia´*** we monitored the effect of differing SSB concentrations on the reaction. Interestingly, our results showed that the inclusion shows a slight increase in the speed of the reaction, whilst not affecting the negative reaction.

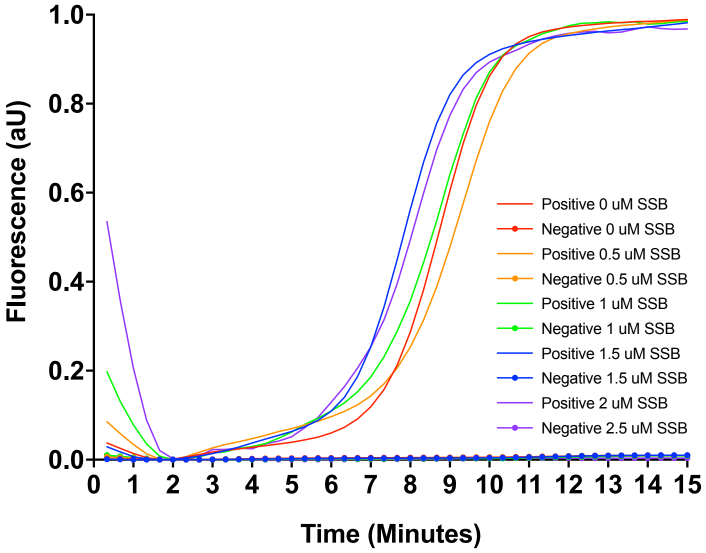


Figure 3.21: *Normalised fluorescence data of SYBR Green I, obtained from EXPAR. (Solid lines) The fluorescence produced from positive samples containing 10 pM of* ***Chlamydia Trigger****, and 50 nM of* ***Template Chlamydia´ – Chlamydia´****. (Dotted lines) The fluorescence produced from negative samples containing no* ***Chlamydia Trigger****, and 50 nM of* ***Template Chlamydia´ – Chlamydia´****. (Red) 0 µM SSB, (Orange) 0.5 µM, (Green) 1 µM, (Blue) 1.5 µM and. (Purple) 2 µM.*

While the inclusion of SSB only provided a small increase on the speed of the reaction, this still prove to be important when it comes to later studies. As a result, we have decided to include 2 µM SSB in all future reactions as this produces significant positive amplification following 15 minutes of incubation with minimal false positive amplification occurring over the same time period.

3.4.5 Reversed-Phase High Performance Liquid Chromatography

3.4.5.1 RP-HPLC of DNA oligomers

Due to the qualitative nature of fluorescent studies, we were only able to indirectly imply successful amplification through an increase in DNA duplex formation and not if ***Chlamydia* Trigger** has been amplified. HPLC enables separation based on the polarity of compounds. This, combined with the ability to continuously monitor the absorbance, enables the identification and quantification of the products of amplification. Through the use of RP-HPLC we designed an elution profile that, when applied to LC-MS would allow for the identification of the masses at each elution. Initial studies were used to determine if two DNA oligomers, with a difference in length of one base, could be separated. The elution times, of **Reporter** (**Figure 3.22**) and ***Chlamydia* Trigger** (**Figure 3.23**) separately and then together (**Figure 3.24**) was analysed. It was possible to see that ***Chlamydia* Trigger** elutes at 16.7 minutes, whereas **Reporter** elutes at 17.5 minutes in all cases. The results indicated that two sequences could successfully be separated. In all chromatograms a rise in absorbance can be seen after 35 minutes, this is a result in the increase of organic solvents washing impurities trapped within the column.

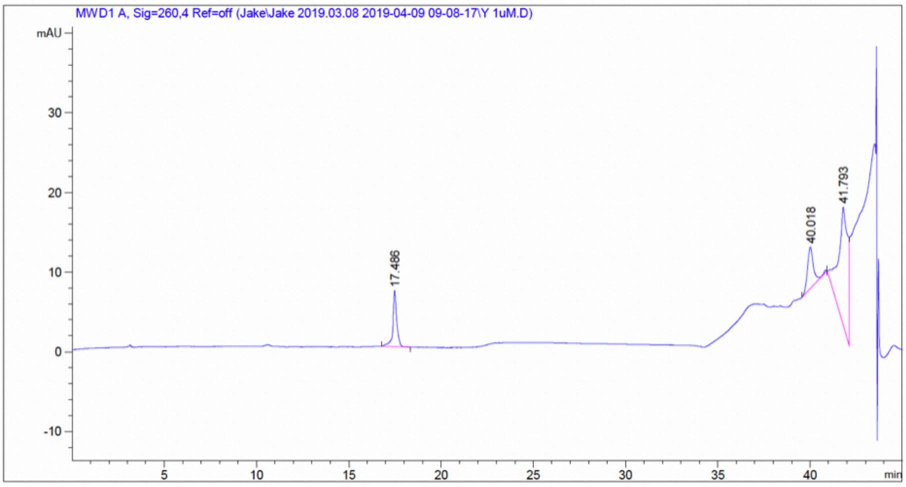


***Figure 3.22:** RP-HPLC Chromatogram generated by **Reporter.** For conditions see Chapter 6.3.7*

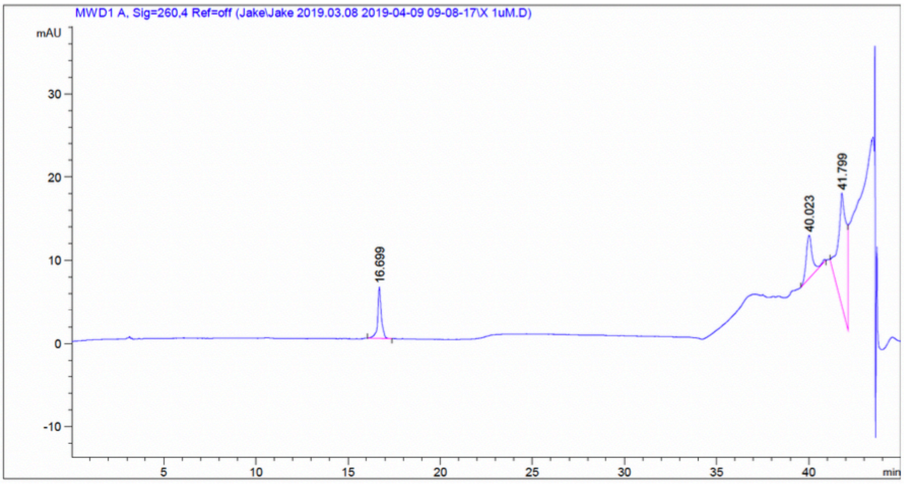


Figure 3.23: *RP-HPLC Chromatogram generated by* ***Chlamydia Trigger.*** *For conditions see Chapter 6.3.7*

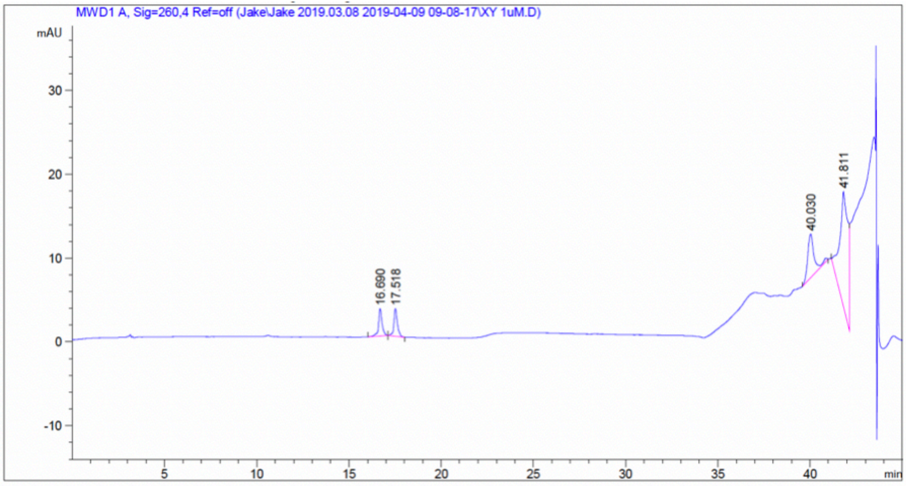


Figure 3.24: *RP-HPLC Chromatogram generated by a mixture containing equal parts* ***Reporter*** *and* ***Chlamydia Trigger.*** *For conditions see Chapter 6.3.7*

3.4.5.2 RP-HPLC of EXPAR

With the elution time of the DNA oligos determined, an appearance of a peak around this time, following incubation of EXPAR for 15 minutes, would imply that oligos have been amplified. RP-HPLC analysis of the EXPAR reaction prior to incubation showed a large peak eluting between 3.7 and 4.3 minutes (**Figure 3.25**) indicating that enzymes and proteins were eluted before the DNA oligomers. Interestingly, **Template *Chlamydia´* – *Chlamydia´*** is not visible, this is likely a result of the concentration being too low to be detected.

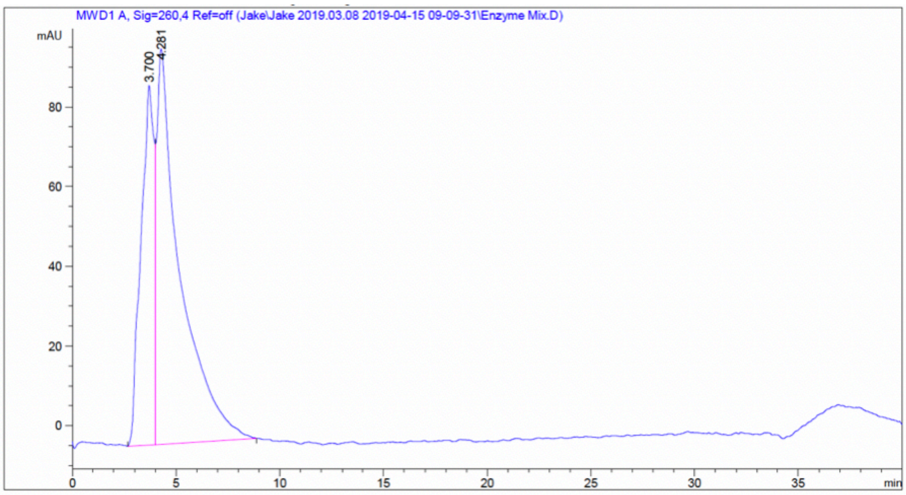


Figure 3.25: *RP-HPLC chromatogram showing the elution of the EXPAR reaction mixture prior to incubation. For conditions see Chapter 6.3.7*

A blank EXPAR reaction was composed using the exponential **Template *Chlamydia´* – *Chlamydia´***, but in the absence of ***Chlamydia* Trigger**. Following incubation at 50 °C for 15 minutes, RP-HPLC was performed (**Figure 3.26**). A broad peak at 4.6 minutes, a result of enzymes and proteins, was eluted followed by a small peak at 20.3 minutes of unknown origin, most likely to be a non-specific amplification product.

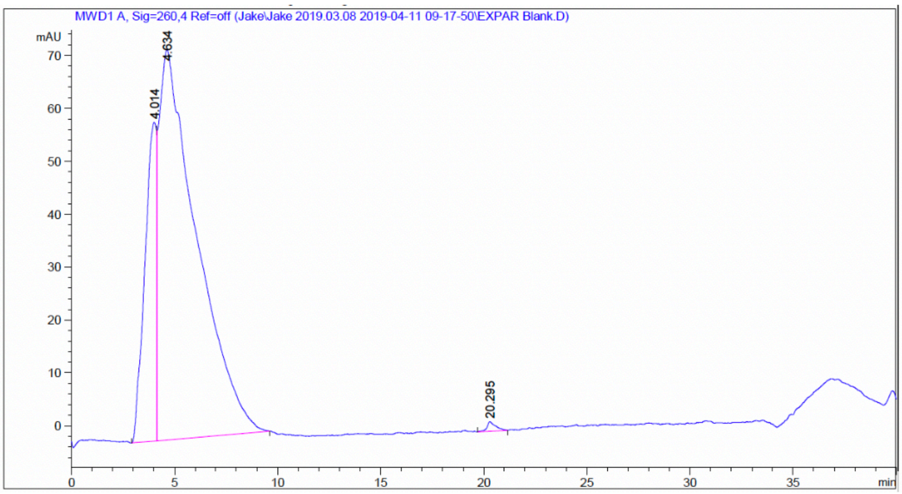


***Figure 3.26**: RP-HPLC chromatogram showing the elution of a negative EXPAR reaction mixture post incubation. For conditions see Chapter 6.3.7*

The inclusion of ***Chlamydia* Trigger** into the EXPAR reaction, thus creating a positive reaction, showed a different chromatogram following incubation. A new peak was eluted at 15.6 minutes, indicating that DNA, most likely ***Chlamydia* Trigger**, has been produced in a quantity significant enough to be detected (**Figure 3.27**). The presence of **Template *Chlamydia´* – *Chlamydia´*** suggests that the product of the reaction should be ***Chlamydia* Trigger** produced from amplification.

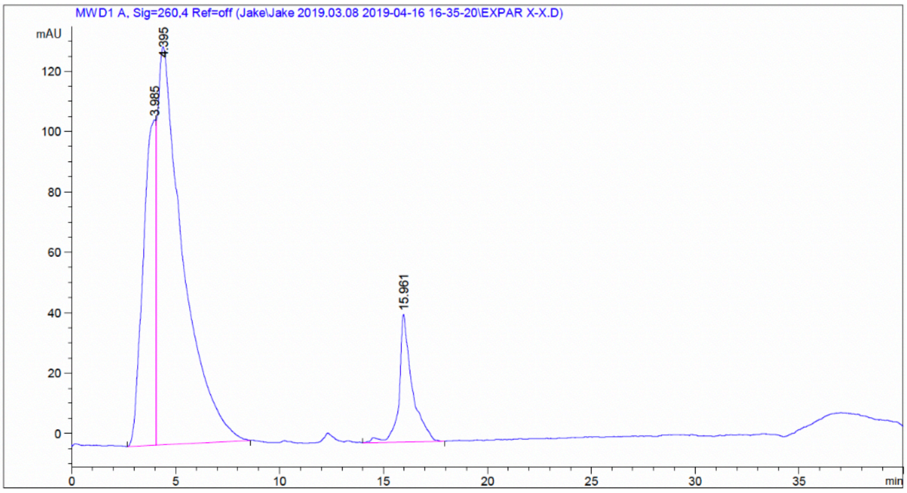


***Figure 3.27**: RP-HPLC chromatogram showing the elution of a positive single stage EXPAR reaction mixture post incubation. For conditions see Chapter 6.3.7*

Further studies were conducted using both **Template *Chlamydia´* – *Chlamydia´*** and **Template *Chlamydia´* – Reporter´** in order to produce the reporter sequence in addition to the trigger sequence. Here we observed the presence of a large addition peak eluting at 14.3 minutes and a smaller peak eluting at 11.7 minutes (**Figure 3.28**). We suspect that the two large peaks at 14.3 and 15.6 are a result of ***Chlamydia* Trigger** and **Reporter**. We believe that the small peak at 11.7 minutes is a result of incomplete strands from unspecific amplification.

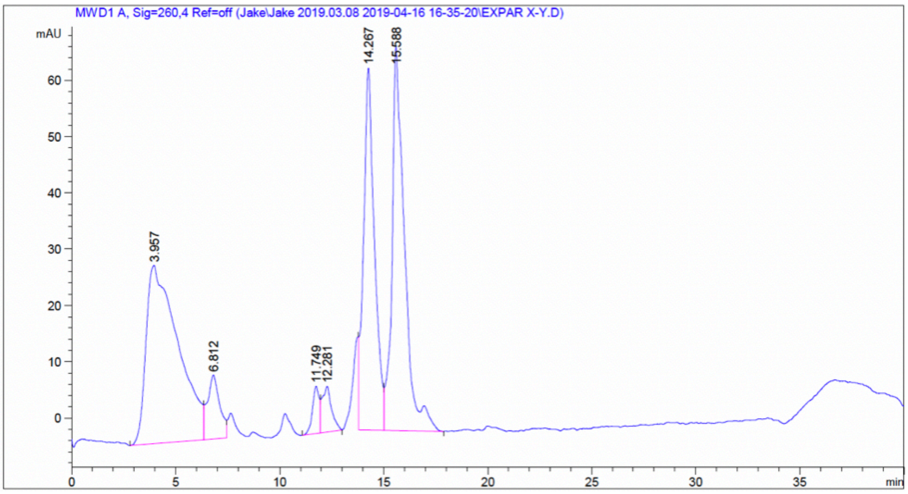


Figure 3.28: *RP-HPLC chromatogram showing the elution of a positive two stage EXPAR reaction mixture post incubation.*

For conditions see Chapter 6.3.7

3.4.6 Mass Spectrometry

As was previously discussed, the product of the EXPAR reaction was modified to produce a sequence with an additional C base at the 5´ end of ***Chlamydia* Trigger**. This additional base has no effect on the amplification; it does however enable differentiation between initial trigger and amplified trigger based on its m/z ratio. Our RP-HPLC studies have provided us with an estimate of where the products of the reaction may elute. The solvent gradient of LC-MS was altered to match that of the RP-HPLC and analysis of an EXPAR reaction containing both templates with and without ***Chlamydia* Trigger** was performed. Analysis of the MS data illustrates that there are no signals produced for the negative control after 8 minutes of incubation. In contrast the positive samples show the presence of amplified ***Chlamydia* Trigger**, due to the presence of two peaks at m/z: 1227.66 (-4) and 1637.2 (-3). As the limit of detection of the mass spectrometer is 1 μM [46], the presence of these two peaks following incubation signifies the reactions has produced in excess of 1 μM of amplified ***Chlamydia* Trigger** from 10 pM of initial ***Chlamydia* Trigger** (**Figure 3.29**). The MS data also enables an assessment to be made of the purity of the products of the EXPAR reaction. The positive samples additionally displayed the presence of two peaks at m/z: 1306.17 (-4) and 1741.58 (-3). These peaks have the same mass as **Reporter** indicating that successful amplification has occurred. In addition to the major peaks two minor peaks, appearing in the same region as **Figure 3.29** relating to ***Chlamydia* Trigger** are also visible (**Figure 3.30**).

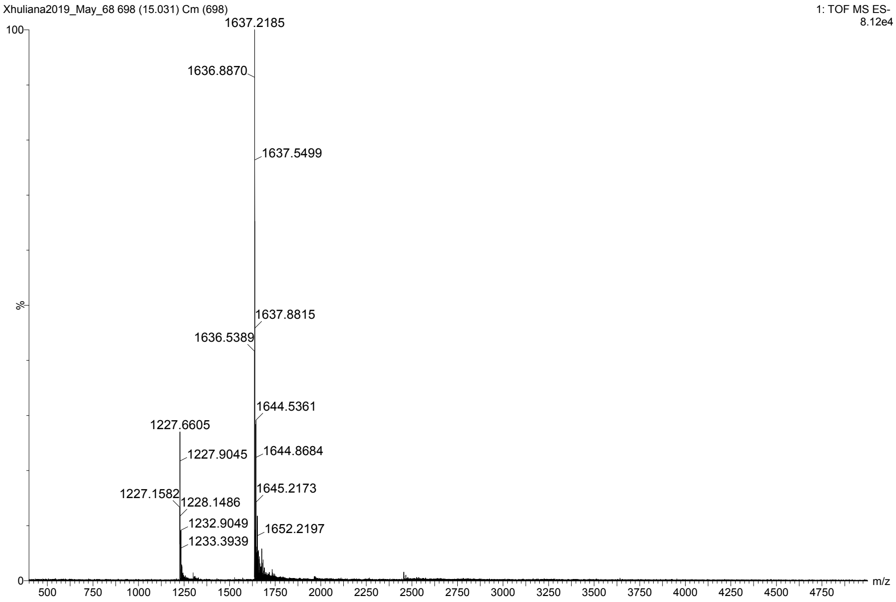


Figure 3.29: *LC-MS data of* ***Chlamydia Trigger*** *produced from a positive EXPAR reaction post incubation (expected m/z: 1227.94 (-4) and 1637.25 (-3)). For conditions see Chapter 6.3.8*

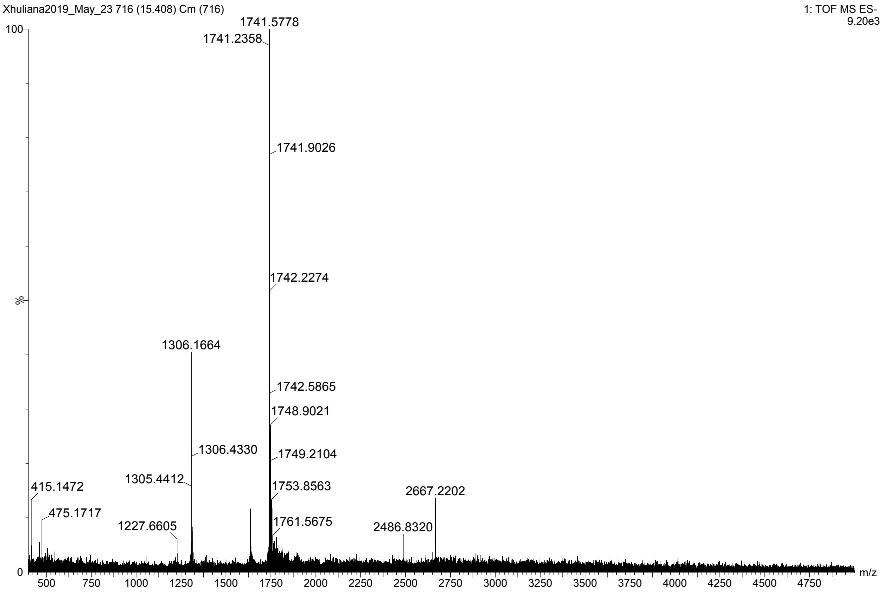


Figure 3.30*: LC-MS data of* ***Reporter*** *produced from a positive EXPAR reaction post incubation (expected m/z: 1306.44 (-4), 1741.92 (-3)). For conditions see Chapter 6.3.8*

3.4.7 Linear Dichroism

We previously discussed the design, synthesis, and testing of M13 Bacteriophage – DNA probes for **Reporter** using LD. As we have established that **Reporter** is produced during the EXPAR process we next needed to combine the two processes. (**Figure 3.12**)

3.4.7.1 Linear dichroic analysis of EXPAR reaction.

Following successful identification of the EXPAR product and production of probes using M13 Bacteriophage – DNA, the next stage was to assess whether the same system could detect **Reporter** produced in the reaction. This has the added difficulty of being carried out in the much more complex EXPAR reaction mixture which includes, enzymes, nucleotides, two templates and two amplification products. The potential for LD to detect EXPAR end products was assessed by performing EXPAR using varying concentrations of ***Chlamydia*** **Trigger** to produce **Reporter**. Following incubation, 10 μL of EXPAR solution was mixed with M13 Bacteriophage – DNA probes. The alignment of the M13 Bacteriophage probe – target complex was then analysed using LD, and a change in LD magnitude was assessed. As already described a decrease in LD signal greater than 5σ of the mean LD signal would prove amplification was successful. In order to assess whether the probes can still be aligned, and the volume of the assay produced an LD reading of 0.01 at 225 nm, a no reaction control measurement containing 10 μL water in place of the EXPAR solution was measured. We were successfully able to see the desired values and LD trace (**Figure 3.*31***).

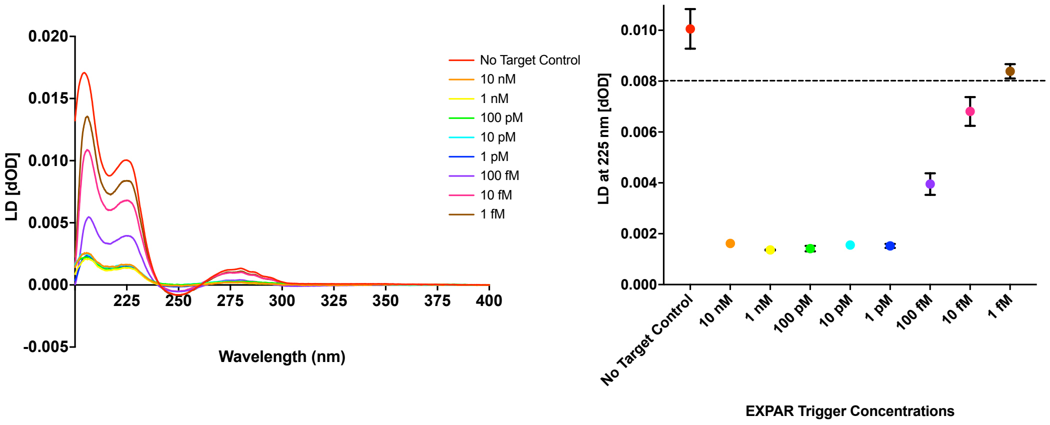


Figure 3.31: *(left) Spectra produced from LD spectroscopy of M13 Bacteriophage – DNA probes containing 10 μL of standard EXPAR reaction solution following incubation with varying* ***Chlamydia Trigger*** *concentrations and their negative controls. (right) LD values of M13 Bacteriophage – DNA probes at 225 nm. Dashed line represents 5σ threshold value. Runs performed in triplicate (n = 3). Error bars in datasets are the standard deviations of the mean.*

The next stage of this study was to assess if the product of the EXPAR reaction could be detected. An EXPAR mixture containing both a linear and exponential template was made, using different concentrations of ***Chlamydia* Trigger**. Following incubation for 15 minutes, a portion of the reaction solution was added to M13 Bacteriophage probes, and the LD spectra was recorded. In **Figure 3.32** we are able to see that the higher concentrations of ***Chlamydia* Trigger** (10 nM – 1 pM) produce a large change in LD signal, thus indicating that a large proportion of the M13 Bacteriophage probes have misaligned. Furthermore, we can see a concentration as low as 10 fM of ***Chlamydia* Trigger** produced enough **Reporter** to cause enough misalignment that the LD signal decreased to below the threshold line.

3.4.7.2 Sensitivity of EXPAR assay.

A decrease in LD signal for samples containing 10 fM or more of ***Chlamydia* Trigger** was observed. While our previous studies of no trigger control reactions displayed little to no amplification over our incubation time frame (**Figure 3.20** & **Figure 3.21**) and no quantifiable products, it is possible that the M13 Bacteriophage probes could detect **Reporter** produced through unspecific amplification if the quantities are greater than 40 nM. In order to assess this, we repeated the same method as previously described (Chapter 3.4.7.1). This time however, we have included no trigger controls, EXPAR reactions that have been incubated in the absence of ***Chlamydia* Trigger**. We performed EXPAR containing 10 fM, 100 fM, 1 pM and 10 pM of ***Chlamydia* Trigger**. Our LD data shows that all four of the positive samples produce a significant decrease in LD signal, whereas all the no trigger controls are still above the threshold (**Figure 3.32**). These studies indicate that the combination of EXPAR with an LD based bioassay creates a rapid isothermal amplification assay with high specificity.

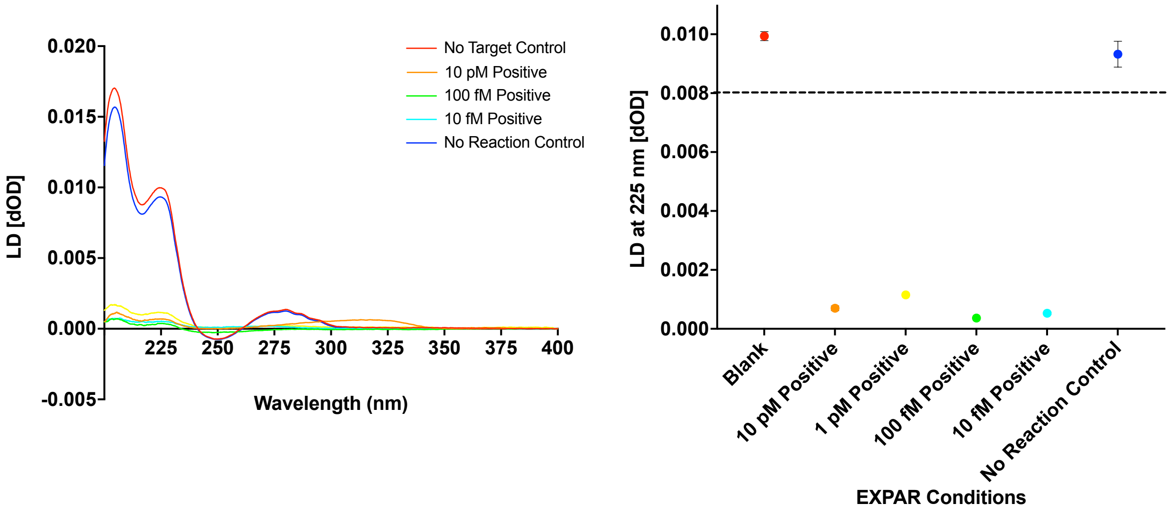


***Figure 3.32**: (a) Spectra produced from LD spectroscopy of M13 Bacteriophage – DNA probes containing 10 μL of standard EXPAR reaction solution following incubation with varying **Chlamydia Trigger** concentrations and their negative controls. (b) LD values of M13 Bacteriophage – DNA probes at 225 nm. Dashed line represents 5σ threshold value. Runs performed in triplicate (n = 3). Error bars in datasets are the standard deviations of the mean.*

3.5 Conclusion

In this work we have identified the successful amplification and detection of synthetic DNA, the first step in the development of an assay. Through improvements to the design of an EXPAR system we have been able to extend the separation time between amplification of positive samples and the unspecific amplification that occurs in the absence of trigger. Following this, we have been able to detect DNA amplified from 10 fM of trigger after an incubation of 10 minutes using the M13 Bacteriophage – LD system. The ability to detect 10 fM of DNA following a 10-minute reaction increases the ability for this system to be developed into a commercially viable point of care assay.

3.6 Future Work

3.6.1 Amplification of Clinical Samples

As discussed, this work focuses on the detection of DNA from a *chlamydia* plasmid, however the DNA used in this chapter was synthetic. Therefore, in order to test the efficiency of the assay it would be necessary to test if the M13 Bacteriophage – probe system is able to detect the amplified product of bacterial plasmid. In order to assess the ability of this system on plasmid DNA, some optimisation would be necessary, for the processing of clinical sample in preparation for analysis. This would enable the sensitivity of the assay system to be compared to other commercially available nucleic acid amplification tests. This is currently being adapted into a working prototype of a *chlamydia* and *gonorrhoea* assay by Linear Diagnostics Limited.

3.7 References

1. Walker, G. T. *et al.* Strand displacement amplification—an isothermal, in vitro DNA amplification technique. *Nucl Acids Res* **20**, 1691–1696 (1992).
2. Spargo, C. A. *et al.* Detection of M. tuberculosis DNA using Thermophilic Strand Displacement Amplification. *Molecular and Cellular Probes* **10**, 247–256 (1996).
3. Walker, G. T. Empirical aspects of strand displacement amplification. *Genome Res.* **3**, 1–6 (1993).
4. Walker, G. T. *et al.* Multiplex strand displacement amplification (SDA) and detection of DNA sequences from Mycobacterium tuberculosis and other mycobacteria. *Nucl Acids Res* **22**, 2670–2677 (1994).
5. Chen, Q. *et al.* Real-time monitoring of the strand displacement amplification (SDA) of human cytomegalovirus by a new SDA-piezoelectric DNA sensor system. *Biosensors and Bioelectronics* **24**, 3412–3418 (2009).
6. Li, J. & Macdonald, J. Advances in isothermal amplification: novel strategies inspired by biological processes. *Biosensors and Bioelectronics* **64**, 196–211 (2015).
7. Craw, P. & Balachandran, W. Isothermal nucleic acid amplification technologies for point-of-care diagnostics: a critical review. *Lab Chip* **12**, 2469 (2012).
8. An, L. *et al.* Characterization of a Thermostable UvrD Helicase and Its Participation in Helicase-dependent Amplification. *J. Biol. Chem.* **280**, 28952–28958 (2005).
9. Jeong, Y.-J., Park, K. & Kim, D.-E. Isothermal DNA amplification in vitro: the helicase-dependent amplification system. *Cell. Mol. Life Sci.* **66**, 3325–3336 (2009).
10. Kivlehan, F., Mavré, F., Talini, L., Limoges, B. & Marchal, D. Real-time electrochemical monitoring of isothermal helicase-dependent amplification of nucleic acids. *Analyst* **136**, 3635 (2011).
11. Vincent, M., Xu, Y. & Kong, H. Helicase-dependent isothermal DNA amplification. *EMBO Rep* **5**, 795–800 (2004).

12. Hall, M. J., Wharam, S. D., Weston, A., Cardy, D. L. N. & Wilson, W. H. Use of Signal-Mediated Amplification of RNA Technology (SMART) to Detect Marine Cyanophage DNA. *BioTechniques* **32**, 604–611 (2002).
13. Maffert, P., Reverchon, S., Nasser, W., Rozand, C. & Abaibou, H. New nucleic acid testing devices to diagnose infectious diseases in resource-limited settings. *Eur J Clin Microbiol Infect Dis* **36**, 1717–1731 (2017).
14. Ali, M. M. *et al.* Rolling circle amplification: a versatile tool for chemical biology, materials science and medicine. *Chemical Society Reviews* **43**, 3324 (2014).
15. Larsson, C. *et al.* In situ genotyping individual DNA molecules by target-primed rolling-circle amplification of padlock probes. *Nature Methods* **1**, 227–232 (2004).
16. Brasino, M. D. & Cha, J. N. Isothermal rolling circle amplification of virus genomes for rapid antigen detection and typing. *The Analyst* **140**, 5138–5144 (2015).
17. Belanger, K. Two-dimensional gel analysis of rolling circle replication in the presence and absence of bacteriophage T4 primase. *Nucleic Acids Research* **24**, 2166–2175 (1996).
18. Gros, M. F., Riele, H. te & Ehrlich, S. D. Rolling circle replication of single-stranded DNA plasmid pC194. *The EMBO Journal* **6**, 3863–3869 (1987).
19. Schweitzer, B. *et al.* Immunoassays with rolling circle DNA amplification: A versatile platform for ultrasensitive antigen detection. *Proceedings of the National Academy of Sciences* **97**, 10113–10119 (2000).
20. Liu, D., Daubendiek, S. L., Zillman, M. A., Ryan, K. & Kool, E. T. Rolling Circle DNA Synthesis: Small Circular Oligonucleotides as Efficient Templates for DNA Polymerases. *Journal of the American Chemical Society* **118**, 1587–1594 (1996).
21. Jeske, H. Barcoding of Plant Viruses with Circular Single-Stranded DNA Based on Rolling Circle Amplification. *Viruses* **10**, 469 (2018).

22. Kuhn, H. Rolling-circle amplification under topological constraints. *Nucleic Acids Research* **30**, 574–580 (2002).
23. Qian, J. *et al.* Sequence dependence of isothermal DNA amplification via EXPAR. *Nucleic Acids Research* **40**, e87–e87 (2012).
24. Mok, E., Wee, E., Wang, Y. & Trau, M. Comprehensive evaluation of molecular enhancers of the isothermal exponential amplification reaction. *Scientific Reports* **6**, (2016).
25. Reid, M. S., Le, X. C. & Zhang, H. Exponential Isothermal Amplification of Nucleic Acids and Assays for Proteins, Cells, Small Molecules, and Enzyme Activities: An EXPAR Example. *Angewandte Chemie International Edition* **57**, 11856–11866 (2018).
26. Asiello, P. J. & Baeumner, A. J. Miniaturized isothermal nucleic acid amplification, a review. *Lab on a Chip* **11**, 1420 (2011).
27. Wang, J. *et al.* Exponential amplification of DNA with very low background using graphene oxide and single-stranded binding protein to suppress non-specific amplification. *Microchimica Acta* **182**, 1095–1101 (2015).
28. Tan, E. *et al.* Isothermal DNA Amplification Coupled with DNA Nanosphere-Based Colorimetric Detection. *Anal. Chem.* **77**, 7984–7992 (2005).
29. Nie, J., Zhang, D.-W., Tie, C., Zhou, Y.-L. & Zhang, X.-X. G-quadruplex based two-stage isothermal exponential amplification reaction for label-free DNA colorimetric detection. *Biosensors and Bioelectronics* **56**, 237–242 (2014).
30. Nie, J. *et al.* Reporter-triggered isothermal exponential amplification strategy in ultrasensitive homogeneous label-free electrochemical nucleic acid biosensing. *Chem. Commun.* **50**, 6211 (2014).
31. Van Ness, J., Van Ness, L. K. & Galas, D. J. Isothermal reactions for the amplification of oligonucleotides. *Proceedings of the National Academy of Sciences* **100**, 4504–4509 (2003).

32. Urtel, G., Van Der Hofstadt, M., Galas, J.-C. & Estevez-Torres, A. rEXPAR: An Isothermal Amplification Scheme That Is Robust to Autocatalytic Parasites. *Biochemistry* **58**, 2675–2681 (2019).
33. Zhao, Y., Chen, F., Li, Q., Wang, L. & Fan, C. Isothermal Amplification of Nucleic Acids. *Chem. Rev.* **115**, 12491–12545 (2015).
34. Ali, A. *et al.* Combining bacteriophage engineering and linear dichroism spectroscopy to produce a DNA hybridisation assay. *RSC Chem. Biol.* 10.1039.D0CB00135J (2020) doi:10.1039/D0CB00135J.
35. Vos, T. *et al.* Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *The Lancet* **388**, 1545–1602 (2016).
36. Harris, S. R., Clarke, I. N., Seth-Smith, H. M., Solomon, A. W. & Cutcliffe, L. T. Chlamydia trachomatis F/SotonF3 plasmid pSotonF3 complete sequence. http://www.ncbi.nlm.nih.gov/nuccore/NC_020988.1 (2017).
37. Harris, S. R. *et al.* Whole-genome analysis of diverse Chlamydia trachomatis strains identifies phylogenetic relationships masked by current clinical typing. *Nat Genet* **44**, 413–419 (2012).
38. Little, H. A. The Development of Novel Diagnostic Sensors Based on Linear Dichroism Spectroscopy. (University of Birmingham, 2016).
39. Carr-Smith, J. *et al.* Polymerase Chain Reaction on a Viral Nanoparticle. *ACS Synthetic Biology* **4**, 1316–1325 (2015).
40. Pacheco-Gómez, R. *et al.* Detection of Pathogenic Bacteria Using a Homogeneous Immunoassay Based on Shear Alignment of Virus Particles and Linear Dichroism. *Analytical Chemistry* **84**, 91–97 (2012).
41. Marrington, R. *et al.* Validation of new microvolume Couette flow linear dichroism cells. *The Analyst* **130**, 1608 (2005).

42. Mori, Y. & Notomi, T. Loop-mediated isothermal amplification (LAMP): a rapid, accurate, and cost-effective diagnostic method for infectious diseases. *Journal of Infection and Chemotherapy* **15**, 62–69 (2009).
43. Tan, E. *et al.* Specific versus Nonspecific Isothermal DNA Amplification through Thermophilic Polymerase and Nicking Enzyme Activities †. *Biochemistry* **47**, 9987–9999 (2008).
44. SYBR® Green Based Quantitative PCR. *Sigma-Aldrich* https://www.sigmaaldrich.com/life-science/molecular-biology/pcr/quantitative-pcr/sybr-green-based-qpcr.html.
45. Real-Time vs Digital vs Traditional PCR - UK. https://www.thermofisher.com/uk/en/home/life-science/pcr/real-time-pcr/real-time-pcr-learning-center/real-time-pcr-basics/real-time-vs-digital-vs-traditional-pcr.html.
46. waters. Xevo G2-XS Q ToF Specification Sheet. *www.waters.com* https://msf.as.ua.edu/wp-content/uploads/2017/08/Xevo-G2-XS-QTof-specification-sheets.pdf.

Chapter 4: Development of a novel assay for the rapid detection of SARS-CoV-2.

4.1 Introduction.

On the 31st of December 2019, 44 cases of pneumonia of unknown etiology were identified in Wuhan City, China.[1] 12 days later it was revealed that the cause was a novel strain of coronavirus related to the Severe Acute Respiratory Syndrome coronavirus (SARS-CoV)[2]. On the 12th of March 2020, 72 days after the first 44 cases were identified, over 20,000 cases had been identified in 118 countries. The World Health Organisation (WHO) declared this new coronavirus as the second pandemic of the 21st century.[3] Scientists worldwide have endeavoured to understand this virus, trying to discover its mechanism of action, effective treatments, develop vaccines and produce faster and more efficient methods of diagnosis to curb its rapid spread.

4.2 Viruses

Viruses are parasitic packages of genetic material (DNA or RNA) wrapped in a protein coat, that while capable of existing outside a host, are unable to reproduce without one.[4] Viruses lack any of the required cellular machinery to replicate and therefore must infect a host and hijack their cells in order to survive and evolve.

4.2.1 SARS-CoV-2

Coronavirus disease 2019 (COVID-19) is an infectious disease caused by Severe Acute Respiratory Syndrome coronavirus 2 (SARS-CoV-2). SARS-CoV-2 is a β-coronavirus, a family of large, enveloped, positive sense, single stranded RNA viruses.[5] There are currently seven known strains of β-coronavirus that are capable of infecting humans,[6] four of which are responsible for approximately 15% of common colds and produce mild symptoms, listed as follows:

1) Human coronavirus OC43 (HCoV-OC43)
2) Human coronavirus HKU1 (HCoV-HKU1)
3) Human coronavirus 229E (HCoV-229E)
4) Human coronavirus NL63 (HCoV-NL63)

The other three strains, however, produce potentially fatal symptoms and are as follows:

1) Middle East Respiratory Syndrome Related Coronavirus (MERS-CoV)
2) Severe Acute Respiratory Syndrome Coronavirus (SARS-CoV)
3) Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2)

Like other β-Coronaviruses strains, SARS-CoV-2 possesses a large RNA viral genome of approximately 30 kilobases capable of coding for 27 proteins of which, 12 play a role in structure and 15 are involved in replication (**Figure 4.1**).[7–11] One of the surface proteins is the spike protein which possesses a receptor-binding domain (RBD). This RBD is fundamental for pathogenic infection as it binds to the human Angiotensin-converting enzyme 2 (ACE2) receptors.[12] Upon the binding of SARS-CoV-2 to ACE2 it is capable of entering cells and beginning its replication cycle, which follows the same basic life cycle as other viral species.[4]

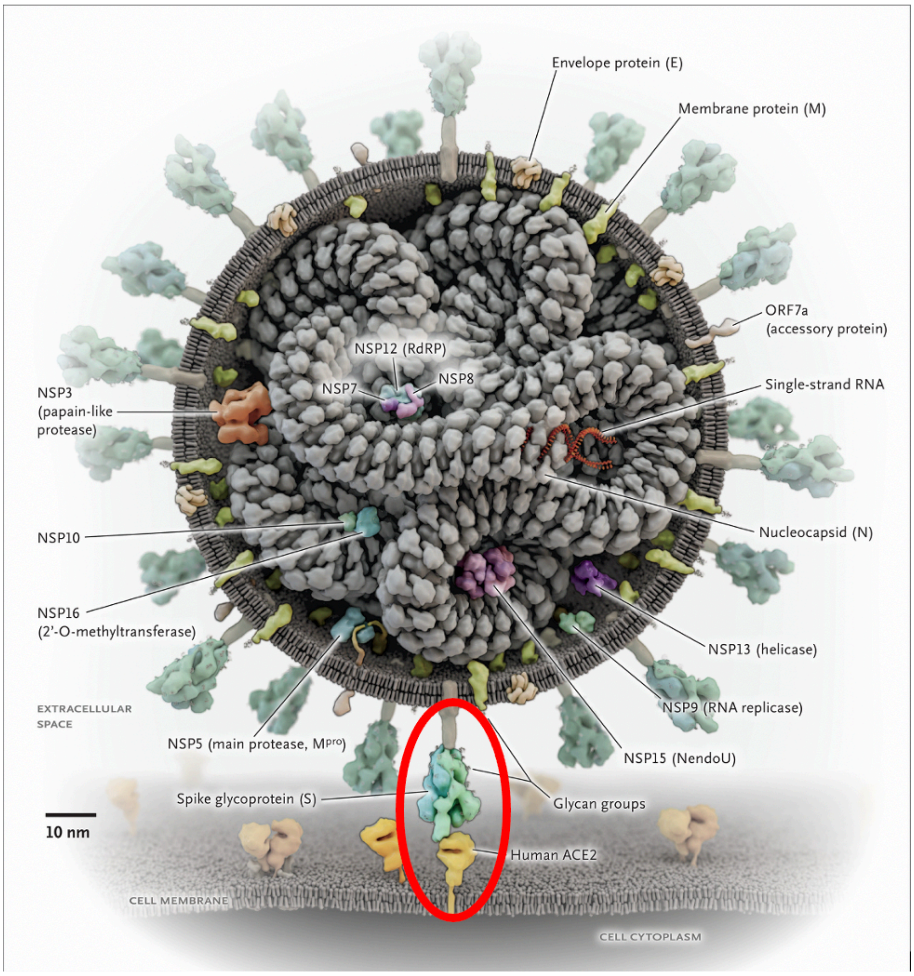


***Figure 4.1**: A diagrammatic depiction of the SARS-CoV-2 virion and its surface proteins. The Spike:ACE2 interaction is circled in red. Image adapted from Parks et al.* [11]

What sets this virus apart from other high-profile diseases of this century such as SARS, MERS, Ebola, and Swine Flu is it is an ideal mix of virulence and lethality compared to some of the most infamous outbreaks of history (**Table 4.1**). This is because SARS-CoV-2 possesses a high basic reproductive rate and a lower-case fatality rate, resulting in a virus capable of spreading through the population while producing a moderately low mortality rate.

***Table 4.1**: A list of infamous diseases, their case fatality rate and estimated basic reproductive number.*

Disease	Case Fatality Rate	Basic Reproductive Number
SARS-CoV-2	1.4%	~2.5[13]
SARS	9.4%	~3[14]
MERS	34%	~0.5[15]
Ebola	50%	~1.8[16]
Swine Flu	<0.1%[17]	~1.5[18]
Spanish Flu	10%	~2[19]
Black Death	60%	~1.4[20]
Smallpox	30%	~5[21]
Measles	2%	~15[22]

4.3 Detection of SARS-CoV-2

Testing is imperative to controlling the outbreak, and this was reiterated by the WHO director general who stated “We have a simple message for all countries: test, test, test. Test every suspected case.”[23] All positive cases of SARS-CoV-2 are logged by the WHO, who have been able to use positive test results to monitor the infections globally. In order to successfully do this, the WHO breaks the world down into 6 regions; The Americas, Europe, Eastern Mediterranean, Africa, South-East Asia, and the Western Pacific (**Figure 4.2**). [24]

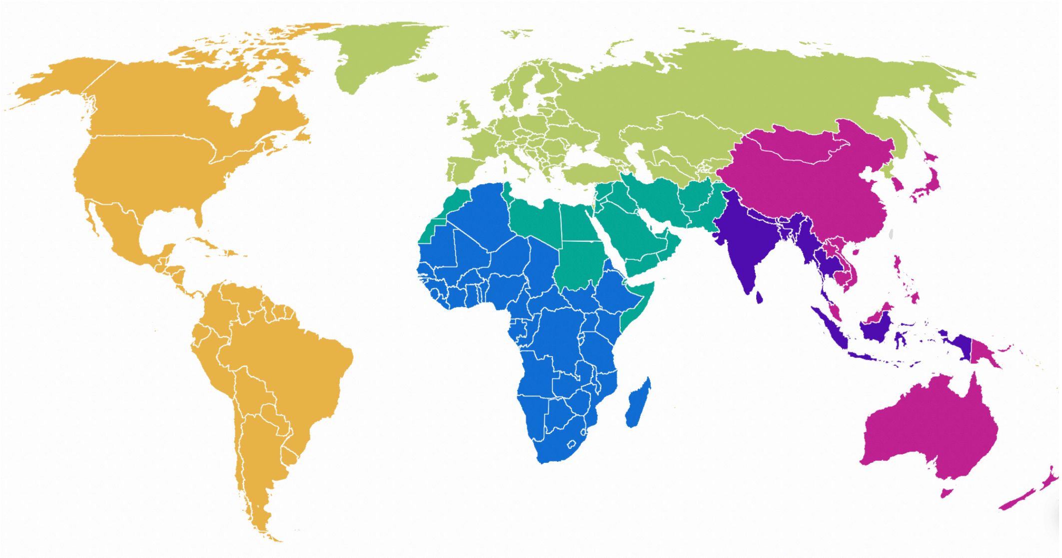


***Figure 4.2**: A map showing the world health organisation regions. (Yellow) The Americas, (Lime) Europe, (Green) Eastern Mediterranean (Blue) Africa, (Purple, South-East Asia, (Pink) Western Pacific.[24]*

Upon identifying a positive case, the WHO are able to log cases from each country into their corresponding regions and then determine the number of cases per week.[25] (**Figure 4.3**) In the 24 months following the initial outbreak the number of cases per week is regularly in the millions. The most commonly used method of testing is a lateral flow test (LFT). While LFTs enable rapid identification of infection, the lack of public trust in positive tests, and the subsequent requirement for nucleic acid amplification tests (NAATs, i.e., PCR) to confirm positivity, slows down the diagnosis time. This then enables further spread of the virus whilst waiting for a positive result to be confirmed. However, the simplicity and ability to put testing in the hands of members of the public has enabled widespread testing to have potentially saved millions of lives. An ideal scenario for an assay would be one that has the speed of an LFT coupled with the assurance of both high sensitivity and specificity, thereby allowing for a simply, yet efficient testing strategy that could be used to reduce, and potentially end, the spread of SARS-CoV-2 worldwide.

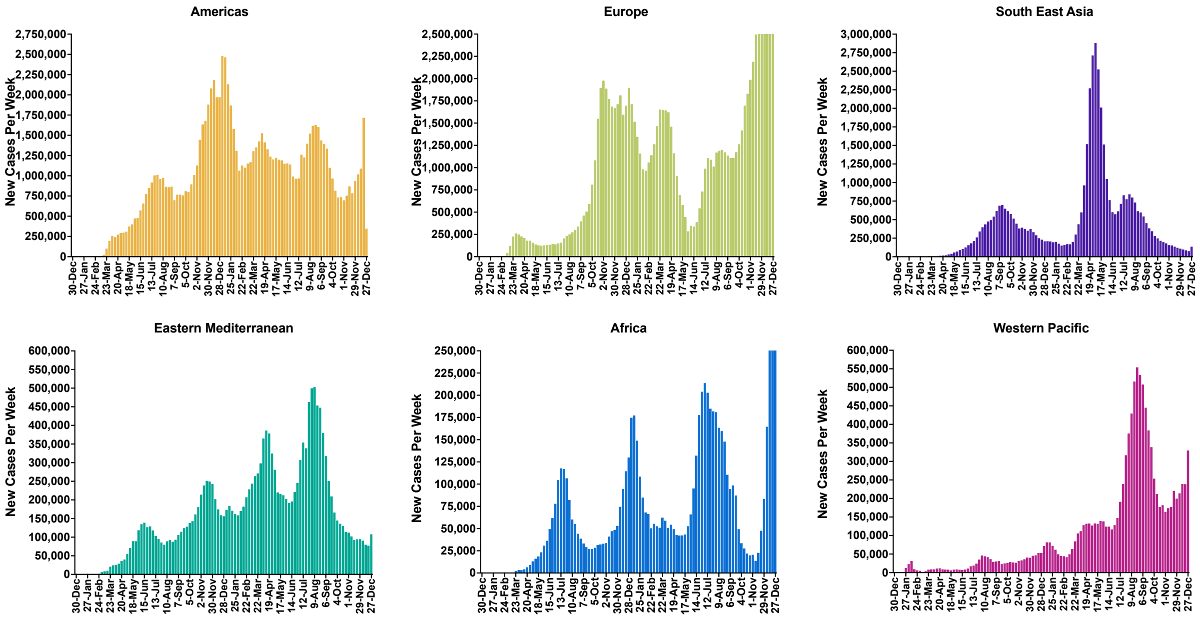


***Figure 4.3**: Graphs depicting the number of weekly cases of SARS-CoV-2 in each of the 6 WHO regions since the first case was identified in December 2019.[25]*

4.3.1 Lateral Flow Antigen Test

LFTs for SARS-CoV-2 (**Figure 4.4**), have been developed to test either for the presence of antibodies in a response to infection (IgG and IgM) or for antigens on the surface of SARS-CoV-2.[9,17,26–30] The advantage of LFTs for the detection of viruses such as SARS-CoV-2 is the ability to detect infectious markers in not only the blood, but also in saliva. This coupled with the fact that no special training of specialists is required, makes LFTs ideal for screening studies in a public setting, and in rural hospitals unlike most other methods of detection.[10,26,28,31–36]

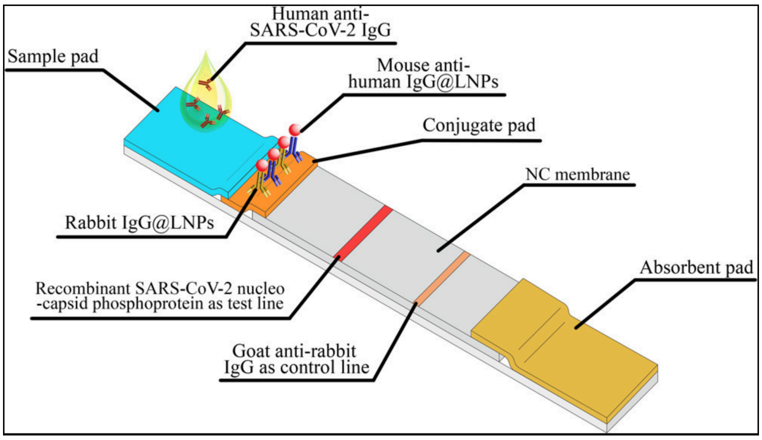


Figure 4.4*: Diagram representing a lateral flow designed for the detection of SARS-CoV-2. Image taken from Chen et al.*[27]

Whilst the use of an LFT for the detection of this virus seems an ideal method there are concerns regarding the accuracy of the test. In a pilot study by the University of Liverpool, samples from asymptomatic patients were compared using lateral flow and RT-PCR. LFTs were comparable with a specificity of 99.68%.[37] However, when the sensitivity was assessed there was a dramatic decrease with LFTs showing a sensitivity of 48.89% in high viral loads, but only 14% when approaching low viral loads.[38,39]

4.3.2 The Challenge of RNA Amplification

There are a plethora of NAATs that are capable of rapid amplification of genetic material. Unfortunately, without modification NAATs are unable to amplify RNA due to the presence of RNAase degrading the structure.[40] In order to overcome this a reverse transcriptase enzyme can be employed[40] to create a copy of the RNA in the form of complementary DNA (cDNA). The reverse transcription (RT) process produces cDNA which can then be amplified through a standard amplification technique.[41] An RT step has been incorporated into numerous assays that are commercially available.[42,43] However, due to the sequential nature of the enzyme the process is time consuming, as longer genomes require more time to be converted into cDNA, as a result this process can take at least 15 – 30 minutes.[44] Following synthesis of cDNA, amplification can begin. The incorporation of RT into the PCR reaction enables the detection of RNA in under 2 hours.[45,46] RT-PCR builds upon PCR (**Figure 4.5**) - briefly described here, a DNA primer anneals to the RNA, where an RT enzyme creates a sequence of cDNA. The temperature of the reaction is raised to denature the duplex and then cooled, allowing further primers to anneal to the cDNA. The mixture is then reheated, and a DNA polymerase creates a full duplex. The heating and cooling process is repeated multiple times creating more sequences of duplex DNA in an exponential fashion.[41] Through the inclusion of a marker, be that an intercalating dye[41] or a beacon/quencher system,[47] amplification can be monitored, and used to assess amplification.

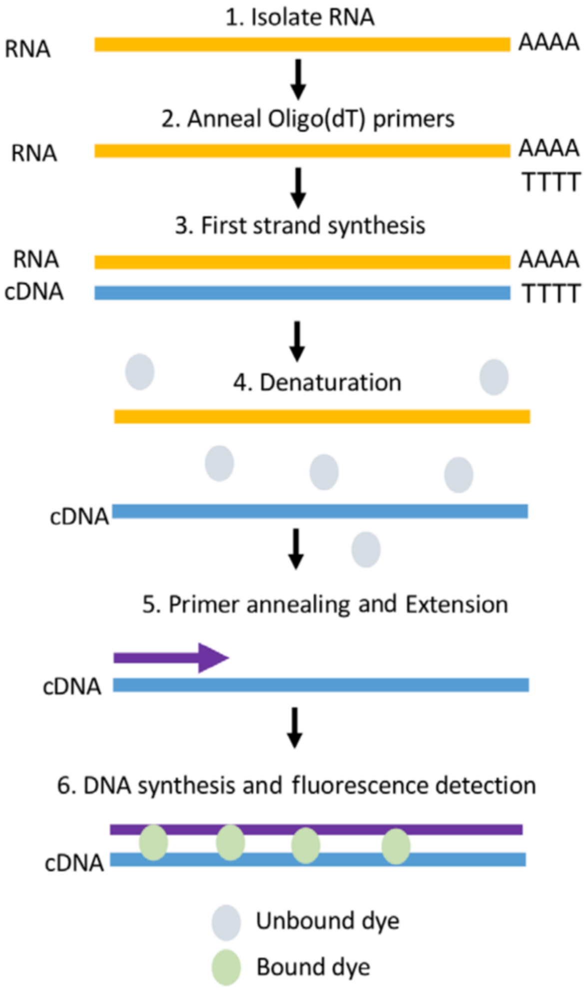


***Figure 4.5**: Depiction of the RT-PCR process, RNA is initially converted into cDNA by reverse transcription. The cDNA is then amplified by PCR. Image adapted from Adamsl.*[41]

RT-PCR as described above is a commercially available test capable of detecting SARS-CoV-2, that has been employed in numerous testing labs, such as the Lighthouse labs.[48] While this is an extremely accurate system to detect the presence of RNA, the requirement of a thermocycler, and the total reaction time, limit the efficiency of RT-PCR. Therefore, significant research has been channelled into developing faster methods of detecting the presence of SARS-CoV-2 RNA. (**Figure 4.6**)

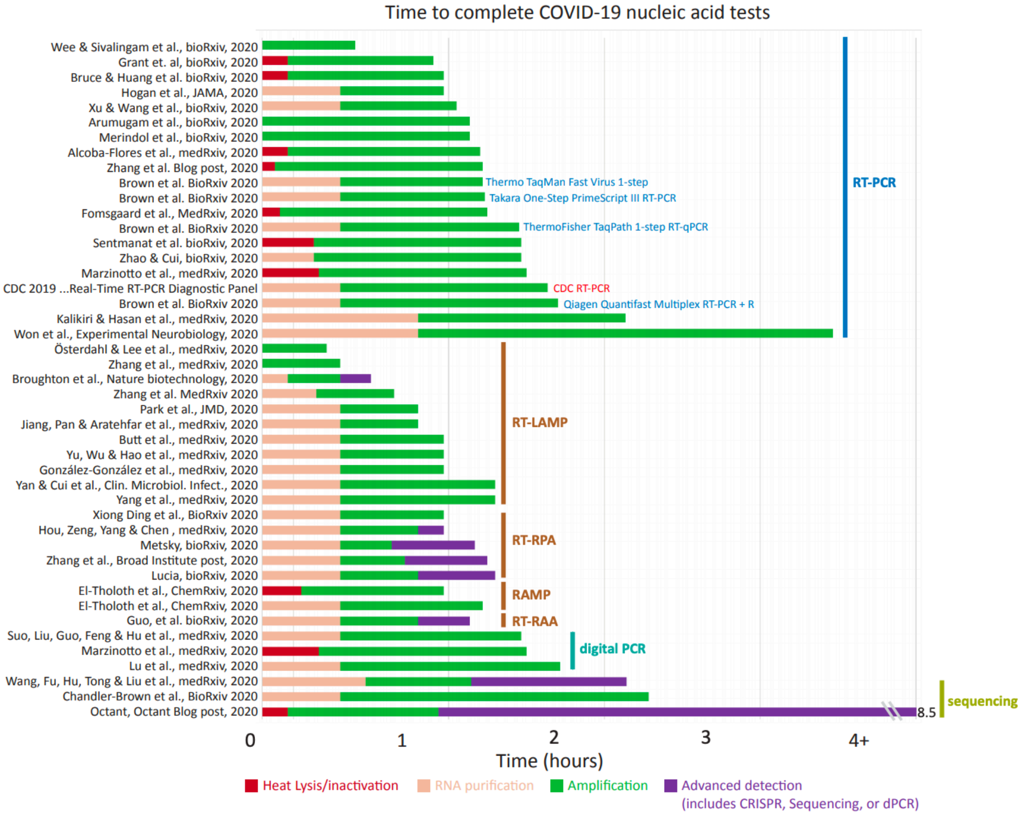


***Figure 4.6**: Snapshot of the time taken for novel assays developed for SARS-CoV-2 to generate a signal. Image taken from Esbin et al.*[49]

Aside from work concerning PCR, a favoured approach is using the isothermal amplification technique loop-mediated isothermal amplification (LAMP). As with PCR, LAMP is unable to directly amplify RNA and therefore an RT step is also required; however, the single temperature approach for LAMP enables reverse transcription and amplification to occur simultaneously, thus decreasing total reaction time from upwards of 90 minutes to 30 minutes, a significant improvement.[30] As with PCR, the inclusion of markers, such as an intercalating dye enables the monitoring of amplification. Interestingly, LAMP can also be detected in two novel ways. Firstly increased turbidity of the reaction solution can occur due to the concentration of amplification products, which results in the solution becoming cloudy as more and more DNA is produced.[50] Secondly, during amplification DNA polymerases liberate protons after the addition of each nucleotide; if the reaction solution is unbuffered, the release of protons gradually decreases the pH of the reaction solution (Chapter 1.4.3.4.2).[51]

4.4 Project Aims

The two main methods of detecting SARS-CoV-2 are lateral flow antigen testing and nucleic acid amplification tests (NAATs such as PCR and LAMP). Lateral flow tests are fast, simple to use with a high specificity, but tend to be less sensitive. NAATs on the other hand are both highly sensitive and specific but require a time-consuming reverse transcription step to convert RNA into DNA prior to amplification:

1) RT-PCR is a slow process taking upwards of 60 minutes.
2) RT-LAMP, while faster is an extremely complicated process.

Therefore, we aimed to produce a test that was as sensitive as a NAAT but as fast and simple to use as a lateral flow test.

Following the successful production of a test a secondary aim was set:

1) To assess the robustness of RTF-EXPAR in the presence of varying other respiratory pathogens.
2) To verify the sensitivity and specificity of RTF-EXPAR on a panel of patient positive and negative samples.

4.5 Results and Discussion

4.5.1 SARS-CoV-2 Assay

4.5.1.1 EXPAR Scheme

It has previously been discussed in this thesis that the exponential amplification reaction (EXPAR) is capable of amplifying short sequences of DNA which can be monitored in real time through the use of an intercalating dye (**Figure 4.7**). However, as the conversion of RNA into cDNA is a lengthy process (taking at least 15 – 30 minutes),[44] no matter how rapid DNA amplification is, the conversion of RNA to DNA will be the rate limiting step of the assay. In order to circumnavigate the requirement of a reverse transcription step, we designed an alternative method to generate the DNA trigger for EXPAR.

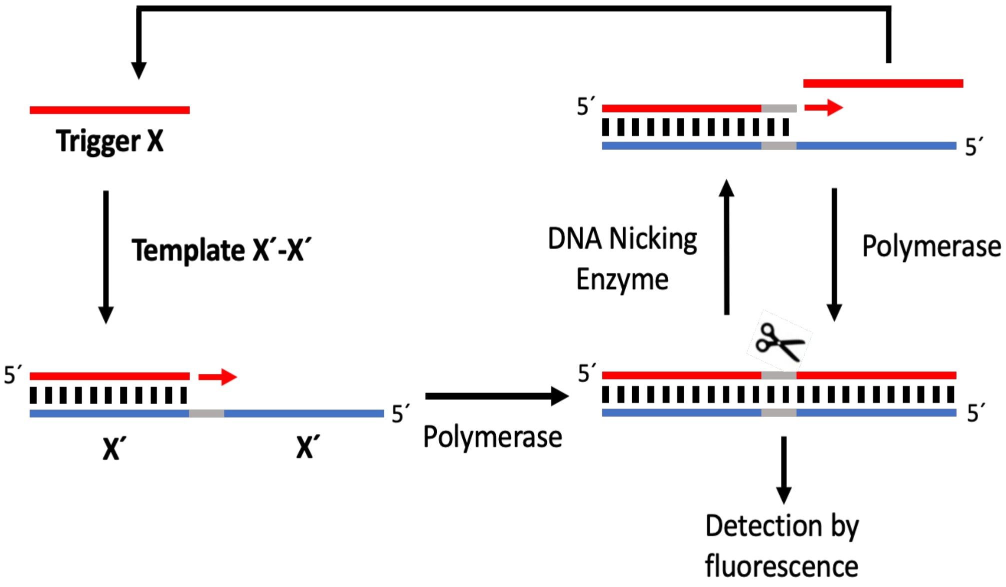


***Figure 4.7**: Schematic of EXPAR. **Trigger X** anneals to **Template X′-X′** and is extended by a DNA polymerase (Bst 2.0 polymerase); the top strand of the newly formed duplex DNA is then cut by a nicking enzyme (Nt.BstNBI); the released DNA (which is displaced by DNA polymerase in a subsequent extension reaction) is identical to **Trigger X** and is therefore able to prime another **Template X′-X′**.*

In a previous study into sequence-specific RNA cleavage by restriction enzymes, Murray *et al* observed that *Bst*NI can selectively cleave DNA within DNA:RNA heteroduplexes.[52] We have taken this finding and used it as a key component to our assay (**Figure 4.10**). We theorised that through forming a heteroduplex between DNA and RNA we would be able to use *Bst*NI to nick the DNA strand of the resulting DNA:RNA heteroduplex to selectivity generate a Trigger for EXPAR. We designed a **Binder DNA** sequence composed of three distinct regions the EXPAR trigger, the *Bst*NI recognition site, and the anchor region, each with their distinct roles in the reaction (**Figure 4.8**).

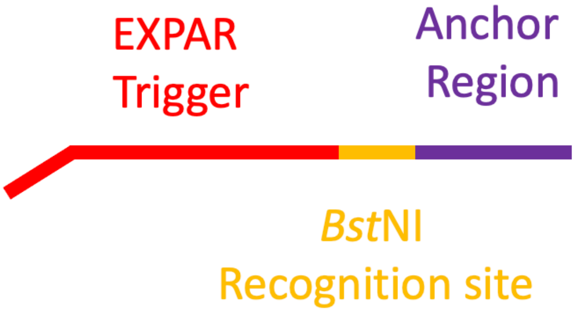


Figure 4.8*:A diagram displaying Binder DNA composed of three regions. Red is the trigger for the EXPAR reaction, this is amplified by EXPAR when RNA is present. Yellow is the recognition site for the enzyme BstNI, this is cleaved when an RNA:DNA heteroduplex is formed. Purple is the Anchor region; this is used to increase specificity of the reaction while preventing the trigger from amplifying when RNA is not present.*

The anchor region is fully complementary to the RNA and increases the T_m of the **Binder DNA** to above the reaction temperature. The recognition site is fully complementary to the RNA and possesses the five bases that enable *Bst*NI to cut the DNA sequence, it should be highlighted that a *Bst*NI site must be presence in the RNA sequence. Finally, there is the EXPAR trigger - this region is partly complementary to the RNA while being fully complementary to the EXPAR template.

The concept behind this design is the **Binder DNA** will anneal to the RNA, with the exception of the non-complementary region (**Figure 4.9**).

Figure 4.9*: When RNA is present, the regions of Binder DNA complementary to the RNA will anneal leaving the non-complementary overhang unbound.*

The annealing of **Binder DNA** with the RNA will create a full duplex over the recognition site, where *Bst*NI will then be able to cleave the DNA of the heteroduplex. This cleavage will result in the EXPAR trigger, and the anchor region being separated and neither region possessing a T_m greater than the reaction temperature, thereby releasing from the RNA. Additional **Binder DNA** will then be capable of binding to the RNA and producing more EXPAR Trigger. The EXPAR Trigger released from the digested mixture is fully complementary to the Template and possessing a T_m close to the reaction temperature, thus enabling amplification to occur (**Figure 4.10**).

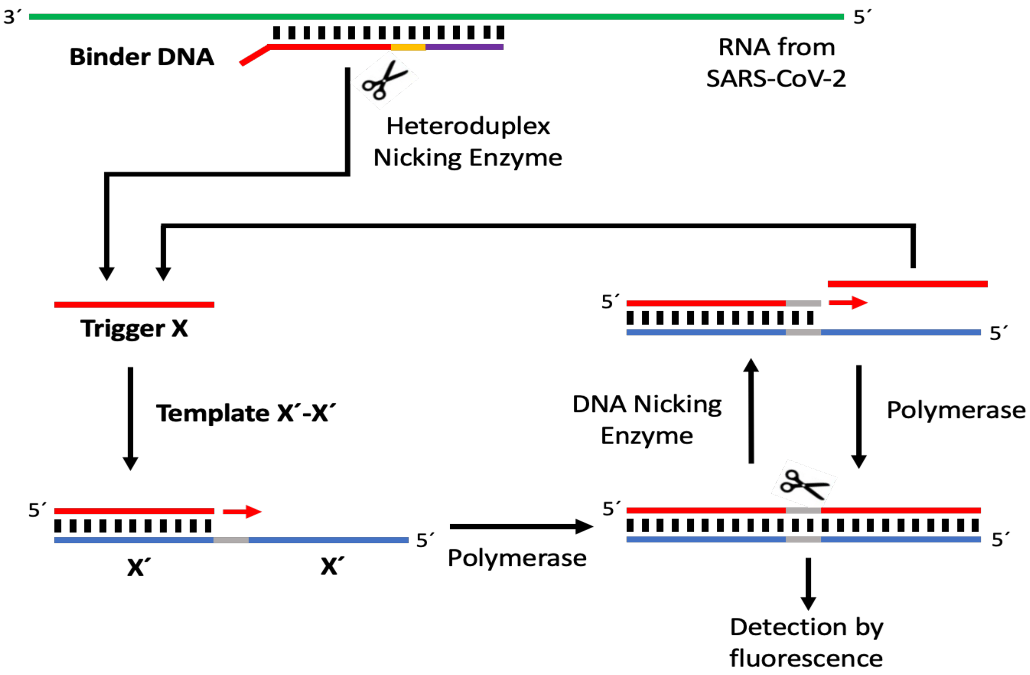


Figure 4.10: *Schematic for Reverse Transcription-Free EXPAR (RTF-EXPAR): Binder DNA anneals to viral RNA; the DNA strand of the DNA:RNA heteroduplex is cut by the restriction endonuclease BstNI, which acts as a nicking enzyme by cutting the DNA strand only, the released DNA strand is* ***Trigger X****, which is then amplified by EXPAR.*[53]

4.5.1.2 Sequence Design

In order for this assay to work, we needed to design three oligonucleotide sequences, **Template X´-X´**, **Trigger X**, and **Binder DNA X.** The most important sequence is **Trigger X** as it is both part of the **Binder DNA X** and complementary to **Template X´-X´**. The recognition site of *Bst*NI[54] is 5´ – CCWGG – 3´ (W being either A or T); therefore we examined the cDNA sequence of the gene *Orf1ab* from SARS-CoV-2 to find every restriction site where the W was a T base, of which there are 10 (**Table 4.2**). T was chosen in place of A in order to limit the number of variables in the initial experimentation. Of these 10 sequences, 4 were unsuitable, three sequences possessed multiple *Bst*NI cleavage sites in close proximity, whereas one possessed a poly A region, thus proving too unstable for a role in EXPAR.

Table 4.2: *cDNA sequences containing the recognition site for the restriction enzyme BstNI with W as an T base (Yellow). The recognition site for the restriction enzyme BstNI with W as an A base (Red).*

Sequence	**Sequence (5´ - 3´)**	**Suitable?**
1	`AAA CCA AAT ACC TGG TGT ATA CGT T`	Yes
2	`TAA ACC ACC GCC TGG AGA TCA ATT T`	Yes
3	`CCT TAA CTT GCC TGG TTG TGA TGG T`	Yes
4	`GAG CCT TGT CCC TGG TTT CAA CGA G`	Yes
5	`CTC ATT CTT ACC TGG TGT TTA TTC T`	Yes
6	`TGT TGA CAT ACC TGG CAT ACC TAA G`	Yes
7	`GCC TGG TTT GCC TGG CAC GAT ATT A`	No, Two Nicking Sites
8	`CCA GGG ACC ACC TGG TAC TGG TAA G`	No, Two Nicking Sites
9	`TTT TGT CGT GCC TGG TTT GCC TGG C`	No, Two Nicking Sites
10	`AAA AAA GGT GCC TGG AAT ATT GGT G`	No, Poly A

Qian *et al* discussed the sequence of the trigger being vital for EXPAR to be successful.[55] The base sequence 5´-`AGGGG–CCAC`-3´ was found to improve EXPAR, with the bases in-between G and C having less of an impact. We have therefore used the findings from this work to design a trigger sequence. The sequence of bases adjacent to the nicking site in EXPAR is important for the reaction to proceed.[55] We therefore designed the 5´ end of the trigger to possess the bases 5´-`AGGGT`-3´, these bases served three roles.

1) To aid in the EXPAR reaction, following the guidelines set forth by Qian *et al*.[55]
2) To act as the unspecific (overhang) region of the Binder DNA
3) To increase the T_m of **Trigger X** on **Template X´-X´.**

Following the inclusion of these bases, the T_m of the **Binder DNA** sequences were analysed (**Table 4.3**). Upon the analysis of the 6 binder sequences, it was possible to ascertain that there was a significant differences between the T_m values of the sequences.[56] Sequence 4 shows a Binder T_m of 61.6 °C, which decreases to 43.5 °C after *Bst*NI cleavage, with the resulting Trigger 4 fully complementary possesses a T_m of 57.2 °C. As the EXPAR reaction was to proceed at 50 °C, the Trigger T_m was deemed too high for EXPAR to proceed as intended and therefore this sequence was discounted. Of the remaining sequences we selected sequence 1, 2 and 3. We decided upon these sequences to monitor a range of Trigger T_m values (1 low (45.6 °C), 3 medium (51.5 °C) and 2 high (55.6 °C)). While sequence 2 possesses a T_m greater than the reaction temperature, it is the only sequence that fits the Qian *et al* sequence criteria,[55] therefore an exception was made. We purchased these three Trigger sequences and their corresponding Templates for our initial studies.

***Table 4.3**: Proposed Binder DNA sequences and theoretical melting temperatures. (Purple) additional AGGGT Sequence, (Yellow) BstNI recognition site, (Green) Sequence of DNA complementary to SARS-CoV-2 RNA after the BstNI recognition site.*

Sequence	**Trigger (5´-3´)**	**T_m (-AGGGT)**	**T_m**	**DNA binder sequence (5´-3´)**	**T_m (-AGGGT)**
1	AGG GTA AAC CAA ATA CC	30.1 °C	45.6 °C	AGG GTA AAC CAA ATA CCT GGT GTA TAC GTT	54.6 °C
2	AGG GTT AAA CCA CCG CC	42.8 °C	55.3 °C	AGG GTT AAA CCA CCG CCT GGA GAT CAA TTT	58.7 °C
3	AGG GTC CTT AAC TTG CC	36.1 °C	51.5 °C	AGG GTC CTT AAC TTG CCT GGT TGT GAT GGT	59.4 °C
4	AGG GTG AGC CTT GTC CC	43.5 °C	57.2 °C	AGG GTG AGC CTT GTC CCT GGT TTC AAC GAG	61.6 °C
5	AGG GTC TCA TTC TTA CC	30.3 °C	47.4 °C	AGG GTC TCA TTC TTA CCT GGT GTT TAT TCT	52.7 °C
6	AGG GTT GTT GAC ATA CC	33.3 °C	48.6 °C	AGG GTT GTT GAC ATA CCT GGC ATA CCT AAG	56.4 °C

4.5.2 Development of an RTF-EXPAR assay

For the development of an RTF-EXPAR assay we were required to complete a two stage process. First, the EXPAR reaction would be developed to determine if the reaction could proceed in the presence of a DNA trigger. Secondly, the development of an RTF step which would feed into the developed EXPAR reaction.

4.5.2.1 Development of an EXPAR assay

The initial studies were used to establish which of the three sequences produced the largest separation between the true amplification and the non-specific amplification (which is always seen with EXPAR as shown in Chapter 3). We performed EXPAR using our established conditions in the presence of 10 nM Trigger (**Figure 4.11**). We determined amplification to be successful when the fluorescence signal was greater than 10 standard deviations away from the variation of the baseline fluorescence. All the samples containing Trigger produced amplification rise times faster than the non-specific amplification. However, we were able to see that Sequence 1 produced the largest separation with a gap between positive and false positive of ~7 minutes, potentially a result of the T_m being below the reaction temperature, therefore destabilising the amplicon. Therefore, we opted to use this as our target sequence.

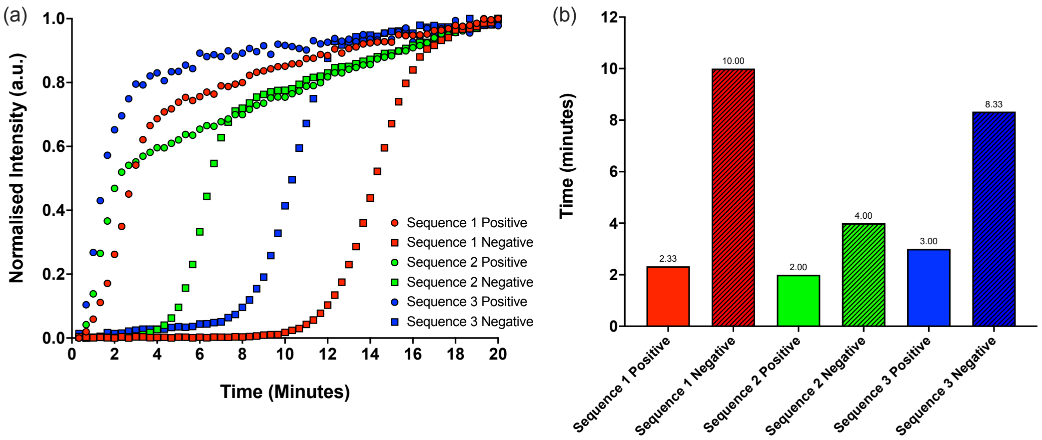


Figure 4.11: *EXPAR sequence test data showing: (a) normalised average fluorescence intensity data (a.u) plotted against time (min) and (b) mean amplification times to greater than 10 standard deviations from the baseline fluorescence, plotted against positive (complementary Trigger at 10 nM) and negative (Water) samples in the presence of templates 1, 2 and 3 (25 nM).*

For the remainder of this work the sequences will be referred to in the following manner (**Table 4.4**):

Table 4.4: *Renamed Trigger sequences.*

Original Sequence Name	New Sequence Name	Trigger (5´-3´)
Sequence 1	**COVID Trigger**	`AGG GTA AAC CAA ATA CC`
Sequence 2	**Trigger SC1**	`AGG GTT AAA CCA CCG CC`
Sequence 3	**Trigger SC2**	`AGG GTC CTT AAC TTG CC`

4.5.2.2 EXPAR Specificity and Sensitivity Determination

With our sequence selected, the next task was to determine the minimum concentration of **COVID Trigger** that would produce an amplification (i.e., a rise time in the fluorescence signal). In order to do this, we performed EXPAR using varying concentrations of **COVID Trigger**. As with the previous tests we determined amplification to be successful when the fluorescence signal was greater than 10 standard deviations away from the variation of the baseline fluorescence. Using the highest concentration of **COVID Trigger** (10 nM), we observed amplification after 3 minutes of incubation, while the no Trigger control produced no amplification until 13 minutes. As the concentration decreased from 10 nM, we observed an amplification time dependent upon trigger concentration. All trigger concentrations of 10 pM and greater showed an increase of amplification signal in less than 10 minutes, showing a distinction from the no Trigger control. Using an unpaired two tailed t-test it is possible to determine if there is 95% confidence that the samples are separated. **COVID Trigger** at 1 pM showed a rise time of 12 minutes, a time that is very close to the no Trigger control, although there is a statistically significant difference ($p = 0.042$). However, **COVID Trigger** 100 fM displayed no significant difference ($p = 0.096$) between amplification of trigger and non-specific amplification from the no Trigger control. This data confirmed that EXPAR is capable of amplifying the SARS-CoV-2 DNA Trigger we designed at concentrations of 1 pM or greater. However, 1 pM was discounted as the time, while statistically different, was not as easy to differentiate compared to 10 pM and above.

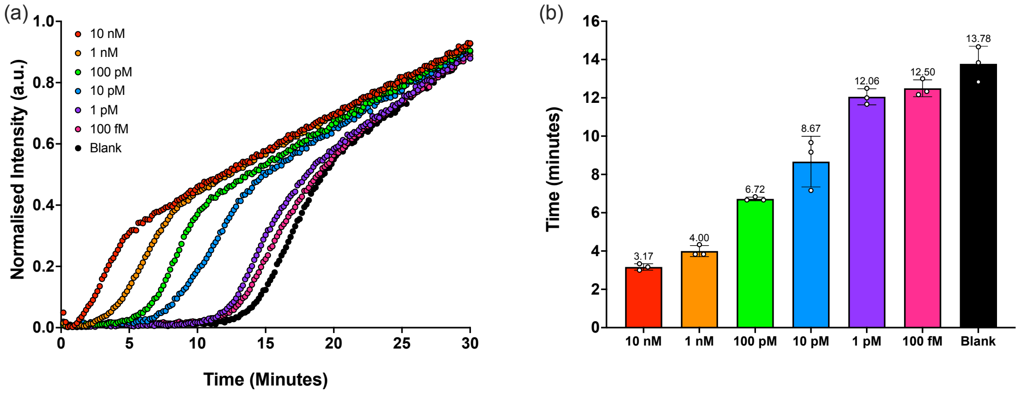


***Figure 4.12**: EXPAR sensitivity test data showing: (a) normalised average fluorescence intensity data (a.u) plotted against time (min) and (b) mean amplification times to greater than 10 standard deviations from the baseline fluorescence, plotted against **COVID Trigger** concentration (10 nM, 1 nM, 100 pM, 10 pM, 1 pM, 100 fM and a no Trigger control) in the presence of **Template COVID´- COVID´** (25 nM). Runs performed in triplicate (n = 3). Error bars in datasets are the standard deviations of the 10-sigma time.*

With the ability of EXPAR to amplify **COVID Trigger** assessed and confirmed, our next study was to test the sequence specificity of the EXPAR process to confirm that other sequences were incapable of triggering amplification at the same speed. We employed the use of the alternate triggers that we designed from the SARS-CoV-2 genome (**Trigger SC1** and **SC2**) and a sequence from *chlamydia* (**Chlamydia Trigger** from Chapter 3) in the EXPAR reaction to determine if non-complementary triggers were capable of priming **Template COVID´-COVID´** and initiating amplification. In order to prevent any non-specific amplification being masked by a low concentration of trigger, we performed each test using 10 nM of trigger, the same quantity that resulted in a 3-minute amplification time for **COVID Trigger**. As shown in **Figure 4.13**, only the fully complementary **COVID Trigger** produced an amplification after 3 minutes, with all the non-specific triggers producing amplification times that were comparable to no trigger at all, with rise times at least three times as long (10 – 13 minutes). This data set confirmed the specificity of the EXPAR reaction, with only the trigger sequence fully complementary to **Template COVID´-COVID´** resulting in fast amplification. With this confirmed, we were able to carry forward **COVID Trigger** for further studies.

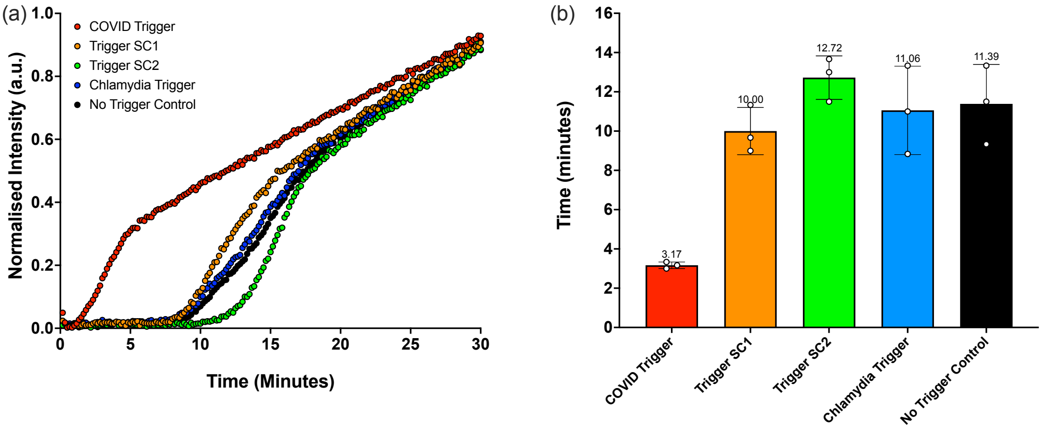


Figure 4.13: *EXPAR specificity test data showing: (a) normalised average fluorescence intensity data (a.u) plotted against time (min) and (b) mean amplification times to greater than 10 standard deviations from the baseline fluorescence signal, plotted against* ***COVID Trigger*** *(fully complementary),* ***Trigger SC1*** *(non-complementary),* ***Trigger SC2*** *(non-complementary),* ***Chlamydia Trigger*** *(non-complementary) and a no Trigger control.* ***Template COVID´-COVID´*** *concentration 25 nM, trigger concentration 10 nM. Runs performed in triplicate (n = 3). Error bars in datasets are the standard deviations of the 10-sigma time.*

4.5.2.3 Development of the reverse transcription free step

To develop a reverse transcription-free step, we were kindly supplied with isolated RNA of SARS-CoV-2, derived from a stock of virus acquired from High Containment Microbiology, PHE, Porton Down. This sample was described as the SARS-CoV-2/human/AUS/VIC16832/2020 isolate, which was originally isolated in Australia from a COVID-19 patient in 2020. The virus was deactivated by heating at 60 °C for 30 min prior to extraction. The concentration was determined by PHE using the Abbott M2000 RT-qPCR Test for SARS-CoV-2 RNA Detection to be 29,080 RNA copies per µL. The samples were subsequently diluted for testing (Chapter 6.4.2).

As previously discussed, *Bst*NI is a commercially available restriction enzyme, that cleaves dsDNA duplex sequences. However, when introduced to a DNA:RNA heteroduplex, *Bst*NI demonstrates selective cleavage of only the DNA sequence.[52] Using *Bst*NI in combination with the **COVID Binder DNA** sequence (containing **COVID Trigger**), we performed a two-pot reaction on SARS-CoV-2 RNA. This involved initially digesting the **COVID Binder DNA**, followed by introducing the digestion mixture into an EXPAR solution, with the intention of seeing amplification in only the digestion solutions containing RNA.

As the concentrations of *Bst*NI and **COVID Binder DNA** were kept constant and in excess, the limiting factor of the reaction would be the quantity of RNA; the theory being the more RNA present, the higher the quantity of **COVID Binder DNA** that can be cleaved and therefore the more **COVID Trigger** that is generated. We performed an enzymatic digestion as adapted from the manufacturer protocol,[54] using 72.5 copies/µL of RNA in the presence of 10 Units of *Bst*NI and 1 µM of **COVID Binder DNA**. The sample was then incubated at 60 °C for 5 minutes. Following the digestion step, 5 µL of the solution was introduced to the EXPAR mix which was then incubated at 50 °C. These results (**Figure 4.14**) showed amplification of the RNA containing sample after 3.17 minutes, with the no Trigger control rising after

10.33 minutes. These data showed it possible to differentiate the sample containing viral RNA from one without. Furthermore, the components of the digestion reaction did not adversely affect the amplification, indicating that a reverse transcription-free (RTF) step is very possible.

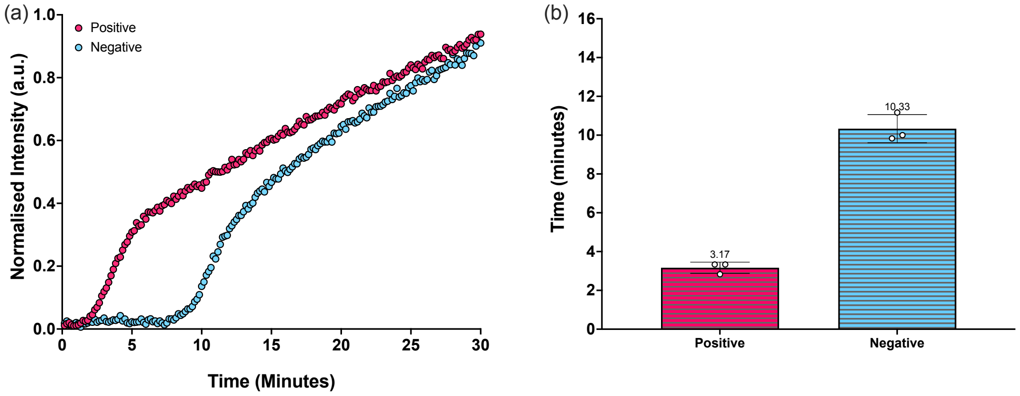


Figure 4.14*: Two-Pot RTF-EXPAR data showing: (a) normalised average fluorescence intensity data (a.u) plotted against time (min) and (b) mean amplification times to greater than 10 standard deviations from the baseline fluorescence signal, plotted against positive SARS-CoV-2 RNA (72.5 copies/µL) and negative (no RNA). Runs performed in triplicate (n = 3). Error bars in datasets are the standard deviations of the 10-sigma time.*

4.5.2.4 Simplifying RTF-EXPAR

While we had determined that a two-pot system for EXPAR detection of RNA was possible, due to the requirement of two reaction solutions and two sets of incubations, the complexity of the assay exceeded our intended aim. Therefore, we endeavoured to test RTF-EXPAR as a one-pot reaction mixture, combining *Bst*NI, **COVID Binder DNA** and the RNA sample into the EXPAR reaction before incubating at 50 °C. While *Bst*NI has an optimal temperature of 60 °C, it functions at 25% efficiency at 37 °C, which suggested that a decrease of only 10 °C would decrease activity but not enough to inhibit function.[54] The results of this study (**Figure 4.15**) demonstrated that RTF-EXPAR could indeed work as a one-pot system, detecting the presence of RNA after 4.00 minutes significantly before the non-specific amplification at 14.78 minutes. While this amplification time is slower than the two-pot system by 0.83 of a minute, the overall assay time was reduced through the combination of the two steps by approximately 5 minutes depending on pipetting and transfer time.

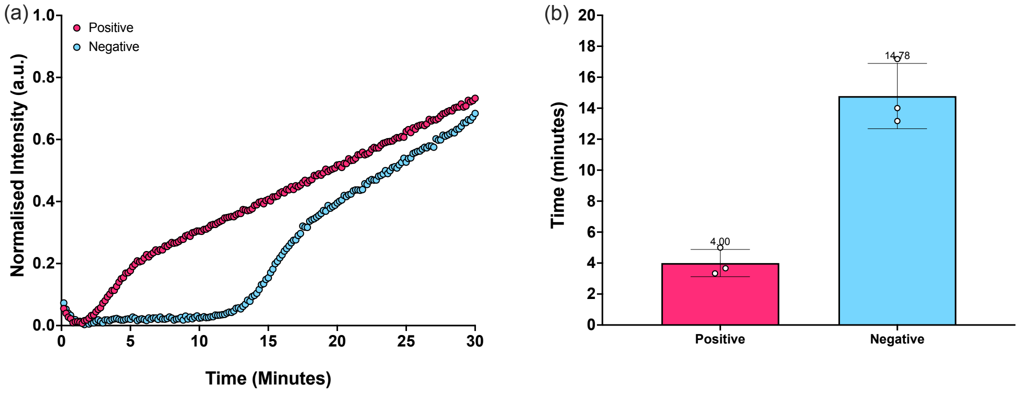


***Figure 4.15**: One-Pot RTF-EXPAR data showing: (a) normalised average fluorescence intensity data (a.u) plotted against time (min) and (b) mean amplification times to greater than 10 standard deviations from the baseline fluorescence signal, plotted against positive SARS-CoV-2 RNA (72.5 copies/µL) and negative (no RNA). Runs performed in triplicate (n = 3). Error bars in datasets are the standard deviations of the 10-sigma time.*

As with the DNA EXPAR studies (**Figure 4.12**), next we next set out to determine how the concentration of sample affected amplification time; to achieve this we introduced a concentration range of RNA into the one-pot RTF-EXPAR mixture prior to incubation. The RNA concentration ranged from 1450 copies/µL to 0.725 copies/µL to simulate a high viral loads (>100 copies/µL), average loads (10 – 100 copies/µL) and low viral loads (<10 copies/µL).[57] As with the DNA samples, we expected the RNA dilution to display a similar trend, with the lower concentrations of RNA displaying a slower amplification time. We incubated each sample at 50 °C in accordance with our previous methodology and were able to see that amplification time was indeed dependant on RNA concentration (**Figure 4.16**). Thus, the highest viral load of 1450 copies/µL produced a rise time 3.08 minutes, comparable with 10 nM of **COVID Trigger**. As the concentration decreases the amplification time increased, with 7.25 copies/µL producing an amplification after 8.75 minutes - this was still statistically significant ($p = 0.012$) before the non-specific amplification occurred. However, this was not the case for 0.75

copies/µL, which did not amplify until 10.50 minutes, which was not statistically different from the no Trigger control ($p = 0.28$). Nevertheless, these results indicated that RTF-EXPAR was capable of detecting low viral loads of extracted RNA.

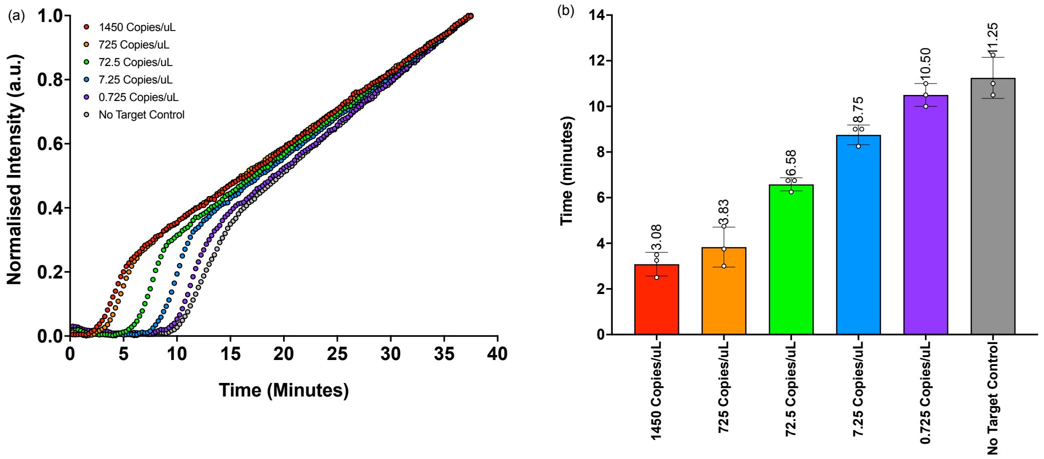


Figure 4.16: *RTF-EXPAR sensitivity test data showing: (a) normalised average fluorescence intensity data (a.u) plotted against time (min) and (b) mean amplification times to greater than 10 standard deviations from the baseline fluorescence, plotted against isolated SARS-CoV-2 RNA concentration (1450 copies/µL, 725 copies/µL, 72.5 copies/µL, 7.25 copies/µL, 0.725 copies/µL and a no Trigger control).*

4.5.3 Verification Studies

The previous sections have outlined the successful design of a novel NAAT for the detection of SARS-CoV-2 RNA without the requirement of a reverse transcription step in under 10 minutes. To assess this assay further, a series of validation studies were now conducted, with the intention of testing the robustness of the assay, the ability of the assay compared to commercially available tests, and to determine the true sensitivity and specificity values.

4.5.3.1 Control data

Initially, we performed a series of control experiments (**Figure 4.17**) to ensure that RTF-EXPAR could only proceed when all the relevant reagents are present. Therefore, RTF-EXPAR was performed using the control conditions listed in **Table 4.5**. The positive control contained all the reagents required for one-pot RTF-EXPAR to proceed, using 72.5 copies/µL of RNA. This was included to ascertain the time taken for an RTF-EXPAR assay containing an average viral load (10 – 100 copies/µL) to produce amplification. An amplification time of 6.58 minutes was consistent with what we had previously seen, again confirming (Chapter 4.5.2.4) that an RTF-EXPAR containing all constituents can, as we expected, successfully produce amplification in the presence of target RNA.

The negative control had exactly the same reagents with the exception of the target RNA, which was replaced with the same volume of water. This test was included to show the time taken for non-specific amplification to occur when all reagents, aside from target RNA, was present. This result was as expected, with non-specific amplification appearing after 11.25 minutes. With the baseline results of a true positive and a non-specific amplification identified, we could compare all other controls to these standards. An ideal outcome of remaining controls should produce amplification times that match or exceed those of the negative control.

***Table 4.5**: Table detailing the reagents of the control studies.*

Control	Water	Buffer	BSA	Bst 2.0 DNA Polymerase	Nt.*Bst*NBI	**Template COVID´-COVID´**	$MgSO_4$	dNTP	SybrGreen	SSB	**COVID Binder DNA**	**SC1 Binder DNA**	*Bst*NI	72.5 copies/µL RNA
Positive Control	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**		**X**	**X**
Negative Control	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**		**X**	
Alternative Binder	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**		**X**	**X**	**X**
No *Bst*NI	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**			**X**
No Binder DNA	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**			**X**	**X**
No Template DNA	**X**	**X**	**X**	**X**	**X**		**X**	**X**	**X**	**X**			**X**	**X**

By replacing **COVID Binder DNA** with an alternative sequence (sequence 9 from **Table 4.3**) we tested the specificity of the binder sequence. In the same way as for the negative control, this reaction mixture was identical to the positive control with the exception that this time the **COVID Binder DNA** sequence was replaced with water. The alternative sequence employed, **SC1 Binder DNA**, is still complementary to a region of SARS-CoV-2. Therefore, *Bst*NI would still be able to selectively cleave **SC1 Binder DNA**, but of course generating **Trigger SC1** (**Table 4.4**) in place of **COVID Trigger**. We have previously tested the effects of **Trigger SC1** in the presence of **Template COVID´-COVID´** which showed only non-specific amplification. Here we saw similar results, with only non-specific amplification occurring after 11.25 minutes, the same time as the negative control, thus confirming that **COVID Binder DNA** is required in order for RTF-EXPAR to proceed efficiently.

We next examined the effects of removing the **Binder DNA** sequence completely, replacing the missing volume with water to negate concentration changes. This test should again show no amplification; the absence of **COVID Binder DNA** means that there is no sequence capable of priming **Template COVID´-COVID´**. As expected, the only amplification observed was a non-specific one timed at 11.50 minutes. This result, combined with that for the alternative binder control, indicates that **COVID Binder DNA** is essential for the RTF-EXPAR assay to function.

Next, through the removal of *Bst*NI, we tested the requirement of the restriction enzyme for the assay. While we knew that the correct sequence of **COVID Binder DNA** is essential for RTF-EXPAR to work, we also needed to ensure that the binder is incapable of generating trigger in the absence of *Bst*NI. As our process is built around the heteroduplex nicking activity of this enzyme, it was vital to the ensure that amplification could not occur without it. As with the other controls, all the remaining components were kept constant, and *Bst*NI was replaced with water. The only amplification observed was at 11.00 minutes, indicative of non-specific amplification.

Our final control was the removal of **Template COVID´-COVID´**, which was also expected to show no amplification. While its removal would not inhibit the formation of **COVID Trigger**, as the **Binder DNA** sequence would still be digested by the enzyme to form a trigger, its amplification would not be possible. As predicted, indeed no amplification was observed after incubating for over 35 minutes, indicating that without the template, no amplification can occur.

The results of our control studies determined that RTF-EXPAR only successfully proceeds in under 10 minutes when all the relevant parts are present - the absence of just one reagent results in non-specific amplification or no amplification at all.

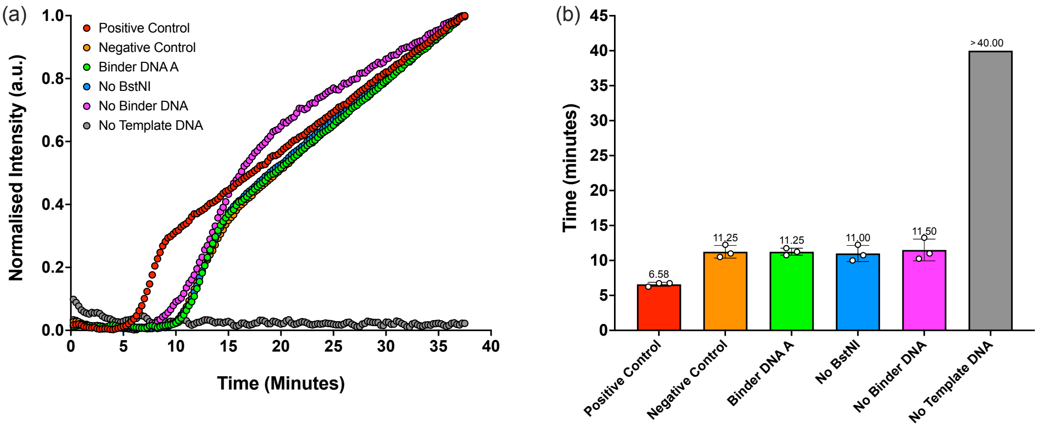


***Figure 4.17**: One-Pot RTF-EXPAR Control data showing: (a) normalised average fluorescence intensity data (a.u) plotted against time (min) and (b) mean amplification time to greater than 10 standard deviations from the baseline fluorescence signal. Data shows positive control (red, 72.5 copies/µL SARS-CoV-2 RNA) containing all components of RTF-EXPAR assay, and five negative controls containing all components of RTF-EXPAR but in the absence of: RNA (orange), Binder DNA X, (replaced by an alternative, Binder DNA A, green), BstNI (blue), Binder DNA (pink) and **Template COVID´-COVID´** (grey). Runs performed in triplicate (n = 3). Error bars in datasets are the standard deviations of the 10-sigma time.*

4.5.3.2 Nucleic acid amplification test comparison

We have determined that RTF-EXPAR is capable of detecting isolated RNA from SARS-CoV-2. In our previous study we determined that RTF-EXPAR is only capable of working when all the constituents are present in the reaction solution. Next, we wanted to test our system compared, both in terms of amplification speed and detection limit, to commercially available assays of the same ilk. RT-LAMP[43] and RT-PCR[42] tests have both been developed for SARS-CoV-2, are commercially available, and are accurate systems. We therefore endeavoured to test the ability of RTF-EXPAR in comparison to these tests. This study was conducted using isolated RNA and the same concentrations range we have previously used in **Figure 4.16**. Each sample was amplified according to standard procedures, be that our newly developed method for RTF-EXPAR or the manufacturers' instructions for RT-LAMP and RT-PCR. The results are displayed in (**Figure 4.18**).

Firstly, we are able to see that RTF-EXPAR is fastest across all samples, while RT-PCR is the slowest. The highest RNA load of 1450 copies/µL produces amplifications after 3.08 minutes for RTF-EXPAR, 11.25 minutes for RT-LAMP and 34.00 minutes for RT-PCR, demonstrating that RTF-EXPAR is ~10 fold faster than RT-PCR and ~4 fold faster than RT-LAMP. At the average viral load range of 72.5 copies/µL again we can see that RTF-EXPAR is the fastest (6.58 minutes) followed by RT-LAMP (13.00 minutes) and finally RT-PCR (38.33 minutes). Low viral loads such as 7.25 copies/µL show amplifications at 8.75 minutes for RTF-EXPAR, 13.83 minutes for RT-LAMP and 39.67 minutes for RT-PCR which shows that RTF-EXPAR is competitive against other tests. As we proceed to lower concentrations still (0.725 copies/µL), we observe amplification times of 10.50 minutes (RTF-EXPAR), 31.92 minutes (RT-LAMP) and 42.67 minutes (RT-PCR). As all NAATs produce non-specific amplification when left long enough, we are able compare the positive amplification times of all three samples assays to their respective no target controls to determine the cut-off point.

RT-PCR displays a non-specific amplification time of 50.67 minutes which is significantly after the lowest RNA concentration of 0.725 copies/µL (p = 0.013) - this is to be expected as it is the most established technique and the 'gold standard' of NAATs. Interestingly, the time for non-specific amplification comes only 1.5 times after the highest RNA concentration. RT-LAMP is faster than RT-PCR in all senses, displaying non-specific amplification after 42.67 minutes, 3.8 times longer than the highest concentration. However, it is an order of magnitude less sensitive than RT-PCR, with the 0.725 copies/µL not being significantly different from the no target control (p = 0.81). RTF-EXPAR is the fastest of all three assays, and the non-specific amplification occurs after 11.25 minutes, 3.6 times longer than the highest concentration of RNA. As with RT-LAMP the sensitivity of our assay is an order of magnitude less than RT-PCR with 0.725 copies/µL showing no significant difference from the no target control (p = 0.28).

In an interesting note, when monitoring the difference in amplification times of positive and negative sample it appears that RT-PCR is presents the largest separation between samples. RT-PCR displays a amplification of positive samples after 34 minutes while the negative samples produced amplification after 50 minutes. However, by looking at the difference in amplification times of the three assays, RT-PCR possesses the smallest percentage difference between non-specific amplification of the negative and the amplification of the highest concentration positive sample with a time difference of 150%. Meanwhile the two isothermal techniques, RT-LAMP and RTF-EXPAR display non-specific amplification times that are respectively 380% and 360% greater than the highest positive.

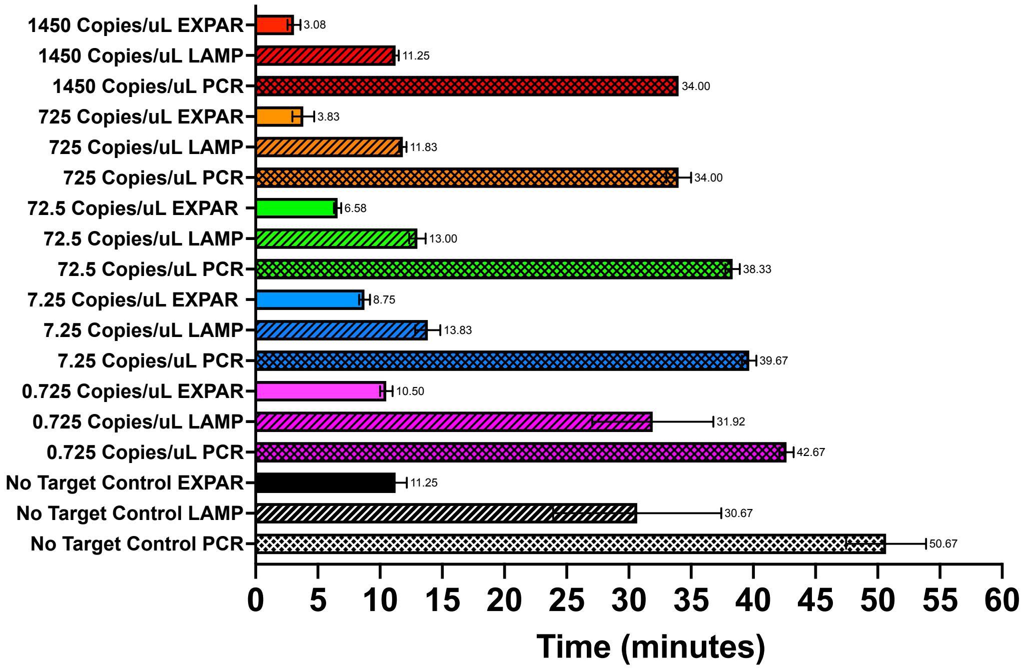


***Figure 4.18**: Three-way comparison data for isolated SARS-CoV-2 RNA detection showing: the mean time for the amplification reaction using RTF-EXPAR, RT-LAMP and RT-PCR. Each run time was calculated to be the point at which the fluorescence signal was greater than 10 standard deviations from the baseline signal (10-sigma time). Error bars in datasets are the standard deviations of the 10-sigma time.*

We have shown that RTF-EXPAR is as sensitive as RT-LAMP, yet faster and simpler than both methods when detecting samples of extracted RNA from SARS-CoV-2. Next, we wanted to test if RTF-EXPAR was capable of detecting SARS-CoV-2 directly from heat denatured virus (**Figure 4.19**). The heat denatured virus was obtained from ATCC. Vero E6 cells infected with SARS-CoV-2 strain 2019-nCoV/USA-WA1/2020 were inactivated by heating. This strain was originally isolated from a human case in Washington state and was deposited by the Centers for Disease Control and Prevention. In this case the RNA of the viral sample was not subjected to an extraction process, only an incubation step at 65 °C in order to kill the virus, thus rendering it safe to handle. The heat denatured samples were diluted from the stock of 420000 genome copies/µL prior to use, then subjected to the same amplification procedures as the isolated RNA. Unlike the isolated RNA used before, this sample is crude, containing viral proteins amongst other impurities and as a result, all three amplification procedures required a longer amplification time before a signal was observed. Interestingly RT-PCR displayed no significant difference between the amplification time of the most concentrated sample (4200 genome copies/µL) and the no target control ($p = 0.66$). This is most likely a result of RT-PCR being optimised to detect pure samples of RNA, suggesting that the introduction of crude material into the reaction solution inhibited the amplification procedure. The RT-LAMP performed much better than the RT-PCR, being able to successfully detect 420 genome copies/µL after 15 minutes of amplification, compared to the no target control of 35 minutes. However, there is a significant drop in performance at the next serial dilution, with no significant difference between 42 genome copies/µL and the no target control ($p = 0.37$). Once again, we believed this to be a result of RT-LAMP being optimised for RNA detection in a purified sample.

On the other hand, RTF-EXPAR performed well with all dilutions from 4200 – 0.42 genome copies/µL having a significant difference from the no target control ($p_{4200} = 0.08$, $p_{420} = 0.02$, $p_{42} = 0.04$, $p_{4.2} = 0.04$ and $p_{0.42} = 0.03$). We have attributed RTF-EXPAR's superior performance to two possible factors. One is that because RT-PCR and RT-LAMP have been highly optimised for pure RNA, crude samples are not suited to their processes. On the other hand, RTF-EXPAR has been optimised to generate the largest gap between true and non-specific amplification, which means that possibly the presence of crude material slows the entire reaction, true amplification, and non-specific amplification alike. A second reason is that RTF-EXPAR functions using a single primer only, compared to the two for RT-PCR or six for RT-LAMP, which is much shorter than the others - this may be beneficial for crude samples, as there is more material in solution that can interfere with the reaction.

Finally, while we may not be able to determine why RTF-EXPAR outperforms RT-LAMP and RT-PCR when using crude samples, it was interesting to see how much better RT-LAMP performed than RT-

PCR. There may be some innate factor in isothermal amplification methodologies that aids amplification of crude samples.

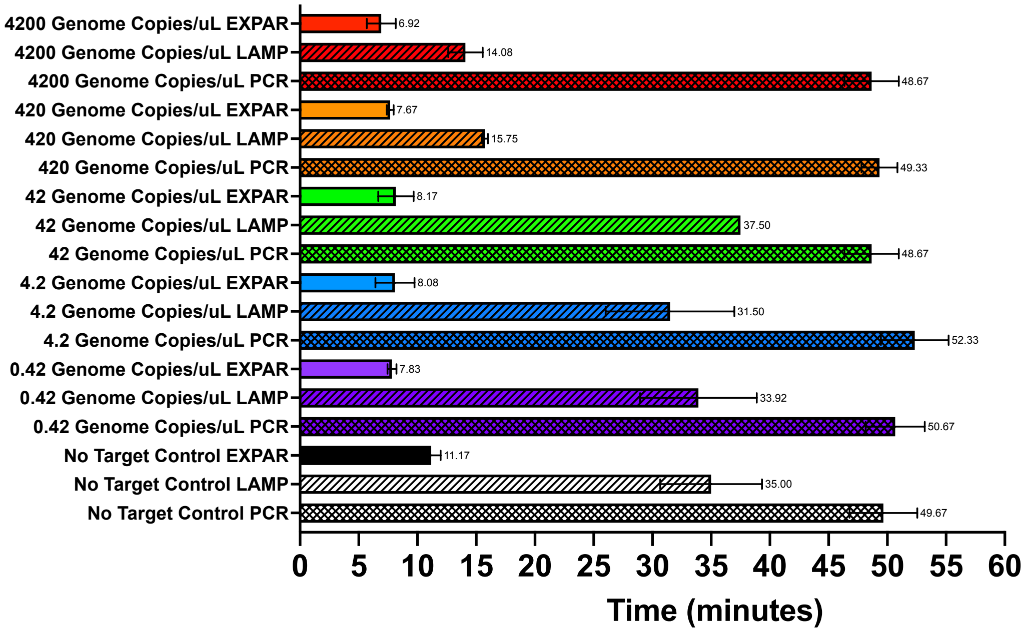


***Figure 4.19**: Three-way comparison data of heat inactivated SARS-CoV-2 detection showing: the mean time for the amplification reaction using RTF-EXPAR, RT-LAMP and RT-PCR. Each run time was calculated to be the point at which the fluorescence signal was greater than 10 standard deviations from the baseline signal (10-sigma time). Error bars in datasets are the standard deviations of the 10-sigma time*

4.5.3.3 Cross reactivity studies.

There are a range of respiratory pathogens which give rise to the same symptoms as SARS-CoV-2. In order to diagnose a patient correctly, a good assay must be able to differentiate between SARS-CoV-2 and other respiratory pathogens (i.e., be specific). Therefore it was decided to test the robustness of RTF-EXPAR in the presence of other pathogenic genomic material using the ZeptoMetrix NATtrol Respiratory Verification Panel 2 (**Figure 4.20**)[58]. In addition to the genomic material for SARS-CoV-2 the panel contained genetic material from the following pathogens: the other six coronavirus species capable of infecting humans (Chapter 4.2) (SARS, MERS, HCoV-HKU-1, HCoV-229E, HCoV-NL63 and HCoV-OC43), four parainfluenza species (HPIV-1, HPIV-2, HPIV-3, and HPIV-4), three species of adenovirus (HAdV-1, HAdV-3 and HAdV-31), Rhinovirus Type 1A, Respiratory Syncytial virus (RSV Type A), Influenza A H1N1 (the cause of the 1918 Spanish Flu, 1977 Russian Flu and 2009 Swine Flu pandemics), Influenza A H3N2 (the cause of the 1968 Hong Kong Flu pandemic), Influenza B, and the bacterial species *Bordetella Pertussis*, *Mycoplasma Pneumoniae* and *Chlamydia Pneumoniae.* Alongside the positive control of SARS-CoV-2 RNA, we also included 10 nM of **COVID Trigger** DNA in order to ensure that the EXPAR reaction itself was functioning as we expected.

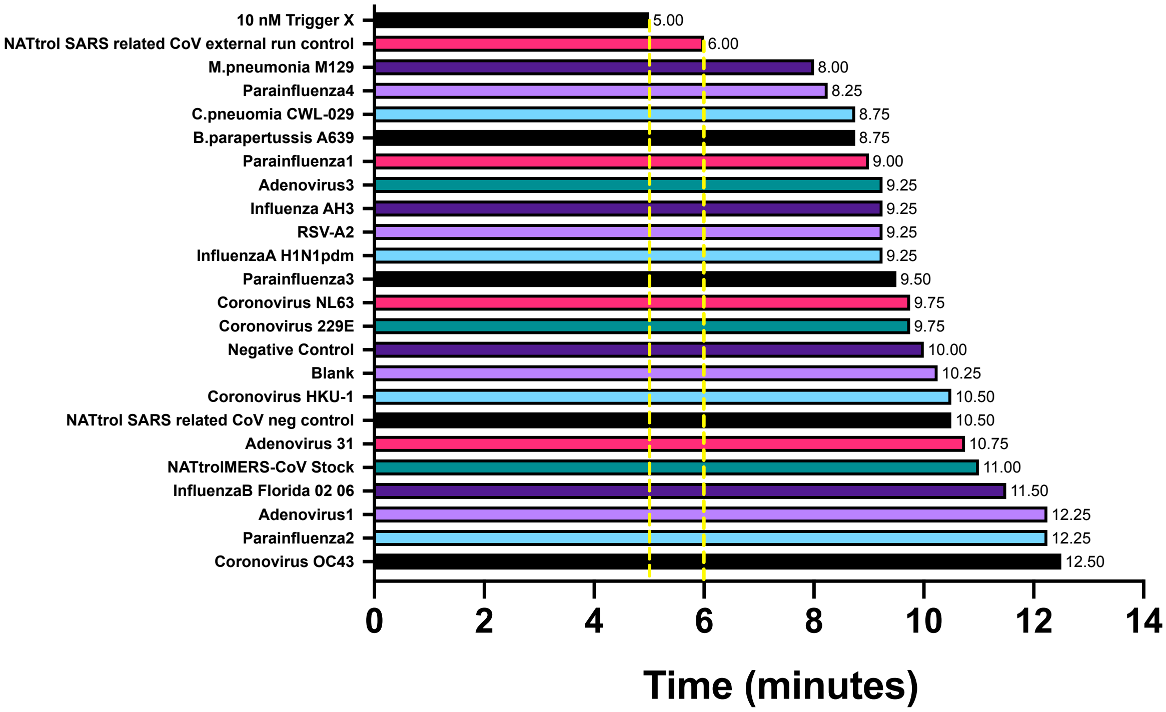


Figure 4.20: *RTF-EXPAR assay data for ZeptoMetrix NATtrol™ Respiratory Verification Panel 2, showing the time for RTF-EXPAR to produce a signal. Yellow dashed lines represent the thresholds for each of the 2 positive controls. Each run time was calculated to be the point at which the fluorescence signal was greater than 10 standard deviations from the baseline signal (10-sigma time). It should be noted that runs against Influenza A H1N1 and Rhinovirus type 1a gave no signal after 40 minutes.*

Analysis of this data shows that our positive controls did amplify the fastest, with **COVID Trigger** amplifying after 5 minutes and the SARS-CoV-2 RNA amplifying after 6 minutes, results which were consistent with what was expected. Furthermore, we were able to see that the no Trigger control reaction produced a rise time of 10.25 minutes and the negative control 10 minutes which were indicative of non-specific amplification. Interestingly, there are 12 species which gave a rise time ranging from 8 minutes – 9.75 minutes, which are all faster than the negative controls, but slower than the true positive. Examination of the nucleotide sequence of these organisms shows that the nicking site for *Bst*NI is present in all of these and at the 5´ end of these sequence there is some sequence similarity. Therefore, it is possible that while not being fully complementary, there is some interaction of the **COVID Binder DNA** strand with other targets that is capable of producing a single **COVID Trigger** which can then go on to prime the EXPAR reaction. Therefore, further optimisation of the conditions (e.g., temperature or concentration) may be necessary. Interestingly, there are also a series of target samples that produce amplification *after* the negative control, including Influenza A H1N1, and Rhinovirus type 1a, neither of which produce any amplification. Overall, there is minimal cross reactivity between the respiratory pathogens and SARS-CoV-2. In the absence of any further optimisation as discussed above, increased amplification speed of some organisms can also simply be managed by adjusting the cut-off point of the reaction to be less than 8 minutes.

4.5.4 Patient sample testing

The final challenge was to assess the robustness of RTF-EXPAR against a range of patient samples graciously supplied by Professor Andrew Beggs from the University of Birmingham. In total there were 276 samples - of these 78 were positive (P) and 198 were negative (N). Initially we could use these values to determine that the prevalence of the virus in the sample population was 28%, as calculated *via* **Equation 4.1**:

$$Prevalence = \frac{Total\ Positive\ Samples}{(Total\ Positive\ Samples + Total\ Negative\ Samples)} \times 100\% = \frac{78}{(78+198)} \times 100\% = 28\%$$

Equation 4.1: *Equation for calculating the prevalence of an infection in a population.*

It should be noted that the 28% prevalence in this study is an increase on the prevalence of SARS-CoV-2 in the community, which at its maximum in England was around 2% at the time of writing.[59] The samples were skewed to have a higher prevalence in order to provide a more in depth analysis of the assay, as a higher positive count will make errors more apparent, thus allowing for determination of a more accurate performance characteristics..

Based upon the data we had collected thus far; we imposed a cut-off time of 6 minutes. Therefore, any samples with a 10 sigma rise time of less than 6 minutes would be deemed to be a positive and any with a time of 6 minutes or greater would be deemed to be negative. The samples were amplified in the same manner as all previous experiments. The results of this study (**Figure 4.21** and **Figure 4.22**) were as follows; 64 true positives (TP), 14 false negatives (FN), 190 true negatives (TN) and 8 false positives (FP). Using the results of this study we can determine vital information about the performance of the assay as outlined below.

It should be noted that all the samples included in this study were certified by PCR, with the 78 positive samples being composed of those that tested positive for either *Orf1ab, N* gene or both. This could have potentially given rise to some positives there were unable to be detected *via* RTF-EXPAR, due to this test only detecting a sequence of RNA in the *Orf1ab* gene.

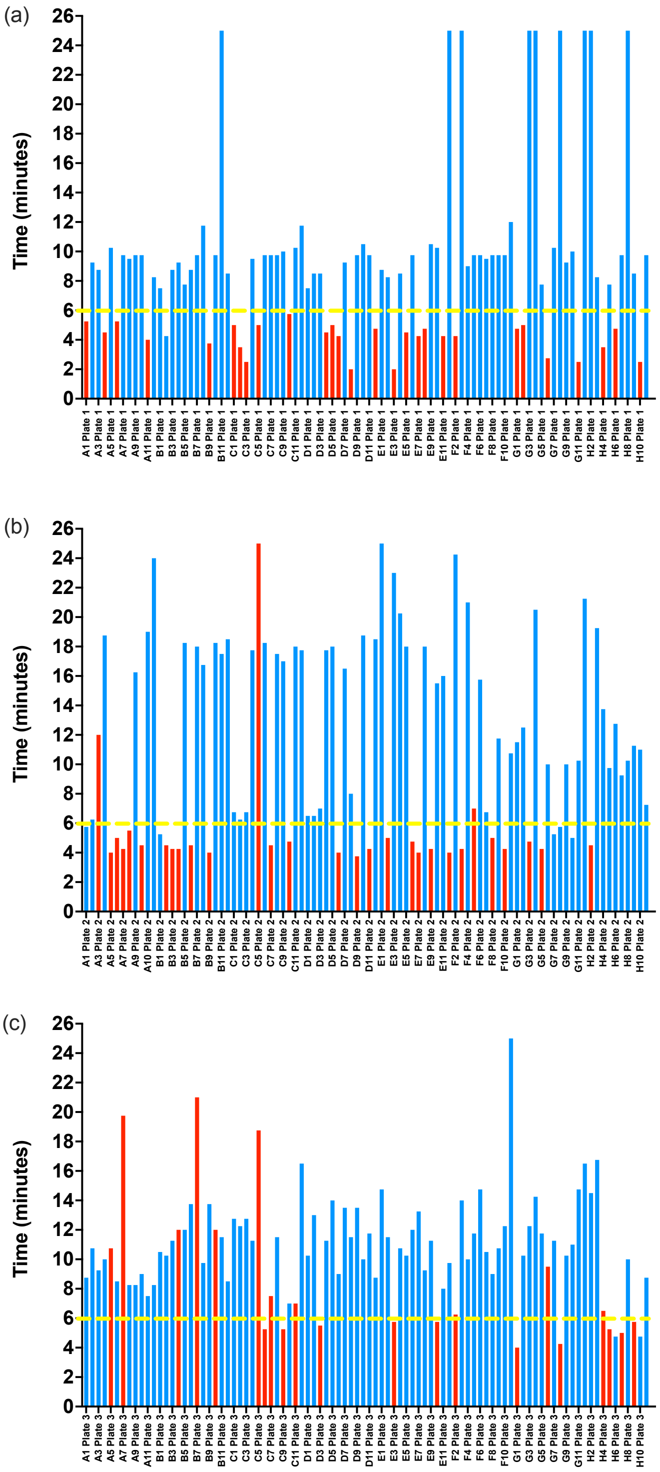
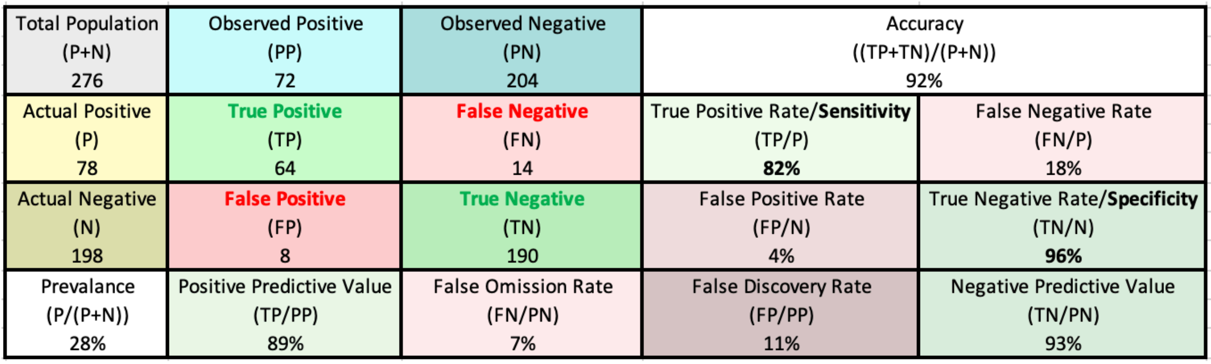


Figure 4.21*: RTF-EXPAR assay data for patient validation samples, showing the time for RTF-EXPAR to produce a signal. Red bars represent positive samples, blue bars represent negative samples, and the yellow dashed lines represent the cut off time of 6 minutes. Each run time was calculated to be the point at which the fluorescence signal was greater than 10 standard deviations from the baseline signal (10-sigma time).*



Total Population (P+N) 276	Observed Positive (PP) 72	Observed Negative (PN) 204	Accuracy ((TP+TN)/(P+N)) 92%	
Actual Positive (P) 78	**True Positive** (TP) 64	**False Negative** (FN) 14	True Positive Rate/**Sensitivity** (TP/P) **82%**	False Negative Rate (FN/P) 18%
Actual Negative (N) 198	**False Positive** (FP) 8	**True Negative** (TN) 190	False Positive Rate (FP/N) 4%	True Negative Rate/**Specificity** (TN/N) **96%**
Prevalance (P/(P+N)) 28%	Positive Predictive Value (TP/PP) 89%	False Omission Rate (FN/PN) 7%	False Discovery Rate (FP/PP) 11%	Negative Predictive Value (TN/PN) 93%

Figure 4.22*: Confusion matrix displaying the results of the patient sample validation study.*

4.5.4.1 Sensitivity

Using this data, we are able to determine the true positivity rate or as it is more commonly known, the sensitivity. Sensitivity is the probability that a test will give a positive result when the patient **has** the illness. Therefore, the higher the sensitivity the more accurate a negative test result will be in determining the patient **does not** have the infection. Sensitivity is calculated using **Equation 4.2**.

$$Sensitivity = \frac{True\ Positive}{Total\ Positive\ Samples} \times 100\% = \frac{64}{78} \times 100\% = 82\%$$

Equation 4.2*: Equation for calculating the sensitivity of an assay.*

The Medicines and Healthcare products Regulatory Agency (MHRA) dictates that in order for a test to be acceptable it should possess a sensitivity of above 80% (against the gold standard of PCR).[60] Using our data set, we were able to determine that RTF-EXPAR possesses a sensitivity of 82%, which is above the recommended level. This means that out of 100 positive samples from 100 patients (as indicated by PCR), 82 give a true positive result by RTF-EXPAR.

4.5.4.1.1 False Negativity Rate

It is also possible to measure the inverse of the true positive rate - this is called the false negativity rate. As the name suggests it is a measure of how many false negatives will be detected from a population of positive samples. This can be calculated *via* **Equation 4.3**

$$False\ negativity\ rate = \frac{False\ Negatives}{Total\ Positive\ Samples} \times 100\% = \frac{14}{78} \times 100\% = 18\%$$

Equation 4.3: *Equation for calculating the false negativity rate of an assay*

Our false negativity rate is 18%, from the 100 positive patients, these are the 18 produce a false negative result, again measured against the PCR result.

4.5.4.2 Specificity

The true negativity rate or specificity, conversely, is the probability that a test will give a negative result when the patient **does not** have the virus. As a result, the higher the specificity, the more accurate a positive test result will be in determining that the patient **does** have the infection. Specificity is calculated using **Equation 4.4**:

$$Specificity = \frac{True\ Negatives}{Total\ Negative\ Samples} \times 100\% = \frac{190}{198} \times 100\% = 96\%$$

Equation 4.4*: Equation for calculating the specificity of an assay.*

The MHRA demands a specificity of 95% for a test to be acceptable.[60] We have determined RTF-EXPAR to possess a specificity of 96%. With both of these criteria met, we have successfully confirmed that RTF-EXPAR has the potential to be an implemented as a testing strategy for SARS-CoV-2.

4.5.4.2.1 False Positivity Rate

As with the true positives, the inverse can also be calculated for the true negativity rate. This takes the name of the false positivity rate and is a measure of the how many negative samples give an erroneous result. This can be calculated using **Equation 4.5**:

$$False\ positivity\ rate = \frac{False\ Positives}{Total\ Negative\ Samples} \times 100\% = \frac{8}{198} \times 100\% = 4\%$$

Equation 4.5*:Equation for calculating the false positivity rate of an assay.*

Due to the high specificity of RTF-EXPAR, the false positivity rate is 4%, indicating that in a cohort of 100 negative patients, only 4 will give an incorrect positive result.

4.5.4.3 Positive Predictive Value

While sensitivity and specificity are the two most important values when it comes to determining an assays efficiency, they are not the only ones. The positive predictive value (PPV), also known as precision, is a measure of the probability that a positive test is truly a positive case. As opposed to sensitivity, which measures true positives against samples which are known to be positive, PPV measures true positives against all samples which test positive in this assay. A higher value indicates that the majority of positive results are true positives, whereas a lower value signifies that the majority of the positive results are false positives and can be calculated using **Equation 4.6**:

$$Positive\ Predictive\ Value = \frac{True\ Positive\ Samples}{True\ Positives+False\ Positives} \times 100\% = \frac{64}{64+8} \times 100\% = 89\%$$

Equation 4.6*: Equation for calculating the positive predictive value of an assay.*

Our test yielded an 89% PPV indicating that for every 100 samples which test positive, 11 will be a result of a false positive.

4.5.4.3.1 False Discovery Rate

The converse of the PPV is the false discovery rate (FDR) which is a measure of the rate of type I errors, or false positive results. This can simply be calculated using **Equation 4.7**

$$False\ discovery\ rate = \frac{False\ Positives}{True\ Positives + False\ Positives} \times 100\% = \frac{8}{64+8} \times 100\% = 11\%$$

Equation 4.7*: Equation for calculating the false discovery rate of an assay.*

As this is the opposite of the PPV, the value is 11%, therefore, 11 people in 100 that will give a false positive result.

4.5.4.4 Negative Predictive Value

The negative predictive value is a measure of the tests ability to correctly identify negative samples, in other words what is the likelihood that a negative result is caused by a negative patient sample. A high negative predictive value indicates that the majority of the samples the tested negative are a result of true negatives, whereas a low value indicated that the majority of samples that tested negative are in fact type II errors, false negatives. We are able to calculate this by using **Equation 4.8**:

$$Negative\ Predictive\ Value = \frac{True\ Negative\ Samples}{False\ Negatives + True\ Negatives} \times 100\% = \frac{190}{14+190} \times 100\% = 93\%$$

Equation 4.8*: Equation for calculating the negative predictive value of an assay.*

RTF-EXPAR gives a negative predictive value of 93%, indicative that of 100 samples which test negative, 93 of these will correctly be from a patient who truly does not have the disease.

4.5.4.4.1 False Omission Rate

As with every other test so far, there is an inverse of the false negative rate, in the form of the false omission rate. A measure of how likely a positive result will be a false negative, and is calculated *via* **Equation 4.9**:

$$False\ omission\ rate = \frac{False\ Negatives}{False\ Negatives + True\ Negatives} \times 100\% = \frac{190}{14+190} \times 100\% = 7\%$$

Equation 4.9: *Equation for calculating the false omission rate of an assay.*

We found RTF-EXPAR to exhibit a false omission rate of 7%, the adverse of the 93% negative predictive value.

4.5.4.5 Accuracy

Finally, we are also able to investigate accuracy. Accuracy is a measure of how many correct diagnoses are made by the test, therefore, how many true positives and true negatives make up all of the results. Evidently, a high accuracy is important, the more accuracy the higher the chances of correctly determining if a patient has or does not have the illness.

$$Accuracy = \frac{(True\ Positive + True\ Negative)}{(Total\ Positive\ Samples + Total\ Negative\ Samples)} \times 100\% = \frac{(64+190)}{(78+198)} \times 100\% = 92\%$$

Equation 4.10: *Equation for calculating the accuracy of an assay.*

Here, we are able to see that RTF-EXPAR possesses an accuracy of 92% indicating that out of 100 samples, 92 will be correctly identified with only 8 being incorrectly identified.

4.5.4.6 *N* Gene Testing

As previously discussed, the majority of NAATs for SARS-CoV-2 target more than one region of the genome. In addition to the *Orf1ab* gene, tests identify the presence of the *N* gene, with a positive signal for either gene indicating the presence of infection. Following the determination that we could develop an RTF-EXPAR system that is capable of detecting a region of the *Orf1ab*, we designed a sequence for targeting the *N* gene of SARS-CoV-2 using the same principles as previously described. Repeating the patient panel studies we were able to see similar results to those shown for *Orf1ab* (**Figure 4.23**).

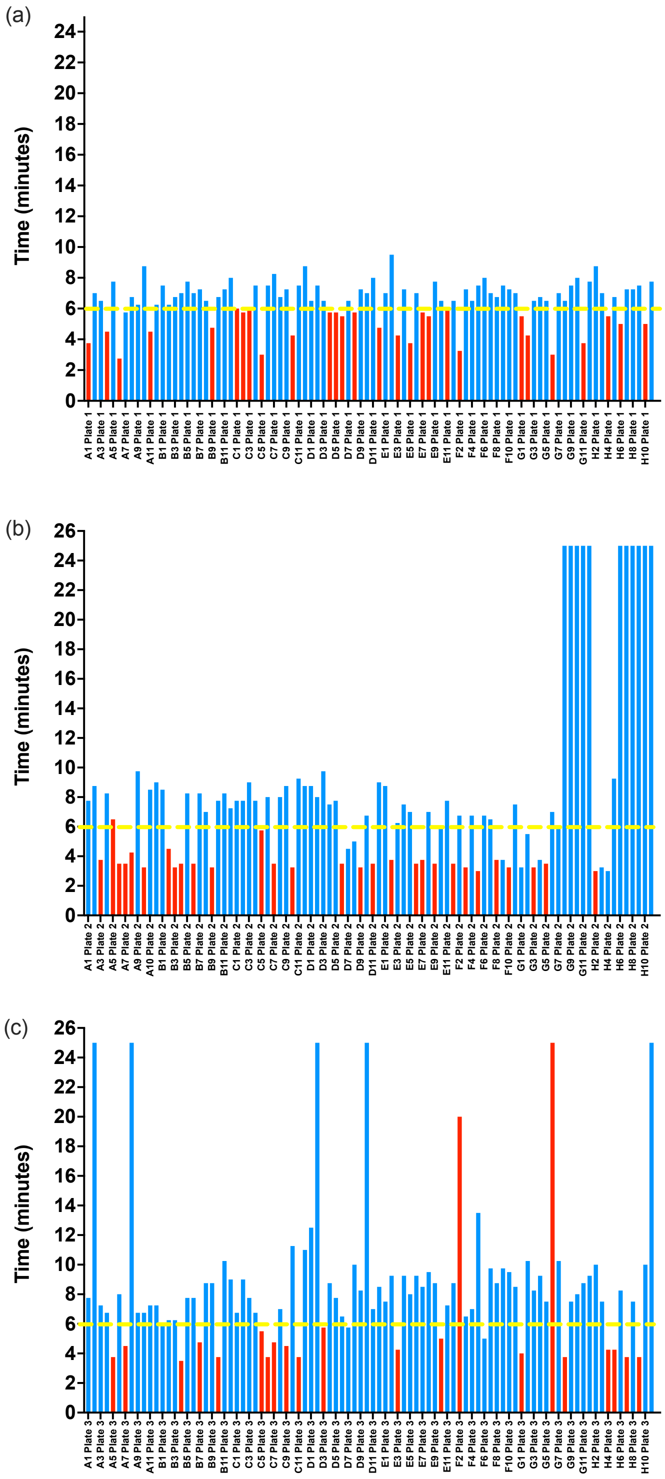
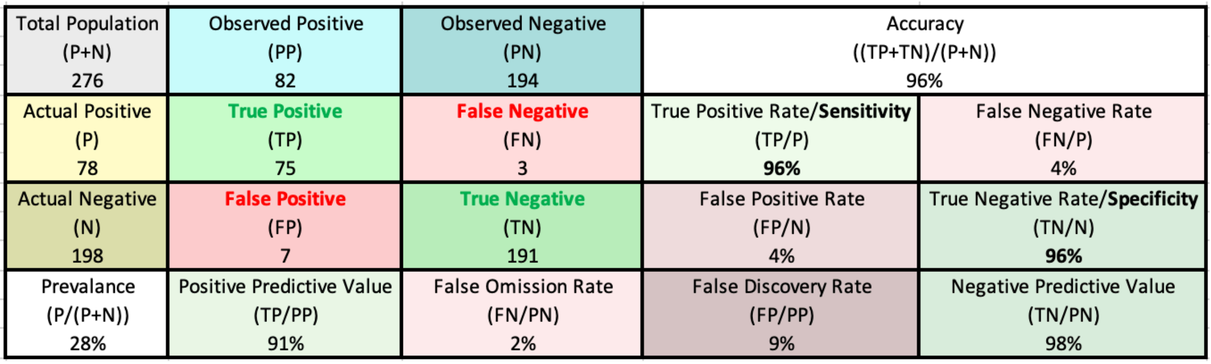


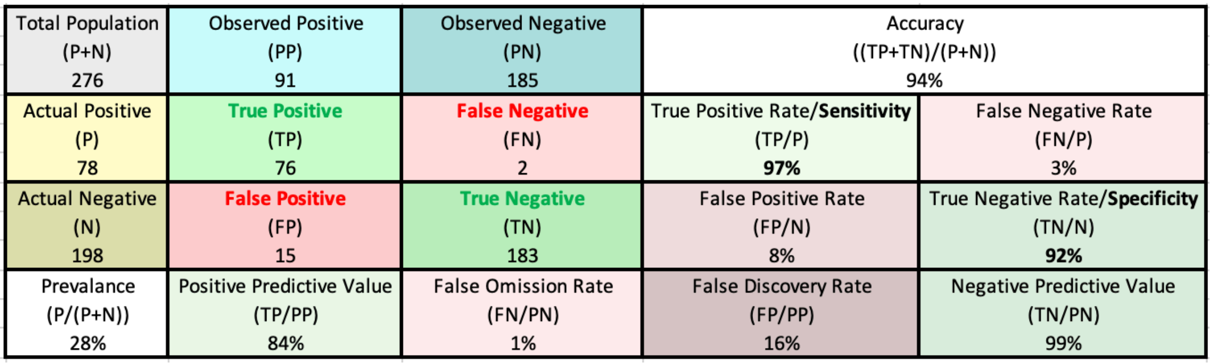
Figure 4.23: *N gene RTF-EXPAR assay data for patient validation samples, showing the time for RTF-EXPAR to produce a signal. Red bars represent positive samples, blue bars represent negative samples, and the yellow dashed lines represent the cut off time of 6 minutes. Each run time was calculated to be the point at which the fluorescence signal was greater than 10 standard deviations from the baseline signal (10-sigma time).*



Total Population (P+N) 276	Observed Positive (PP) 82	Observed Negative (PN) 194	Accuracy ((TP+TN)/(P+N)) 96%	
Actual Positive (P) 78	**True Positive** (TP) 75	**False Negative** (FN) 3	True Positive Rate/**Sensitivity** (TP/P) **96%**	False Negative Rate (FN/P) 4%
Actual Negative (N) 198	**False Positive** (FP) 7	**True Negative** (TN) 191	False Positive Rate (FP/N) 4%	True Negative Rate/**Specificity** (TN/N) **96%**
Prevalance (P/(P+N)) 28%	Positive Predictive Value (TP/PP) 91%	False Omission Rate (FN/PN) 2%	False Discovery Rate (FP/PP) 9%	Negative Predictive Value (TN/PN) 98%

***Figure 4.24**: Confusion matrix displaying the results of the N gene patient sample validation study*

Following the analysis of the *N* gene studies we were able to see similar results to *Orf1ab*, however it is worth noting that there is an increase in sensitivity from 82% to 96%. This increase in sensitivity is a result of the patient samples initially being analysed by RT-PCR for the presence of *Orf1ab* and *N* gene; as a result, samples which tested positive for *N* only would not be picked up by *Orf1ab* RTF-EXPAR. With the results of *N* Gene along with *Orf1ab*, it was possible to combine the results to emulate a standardised test. Any sample that produced an amplification time of less than 6 minutes for either *N* or *Orf1ab* was determined to be a positive test result (**Figure 4.25**). It is clear to see that the sensitivity has increased to 97%, while the specificity has decreased to 92%. These results indicate that testing with two genes the test displays a reduced number of false negatives. Further optimisation can be used to reduce the false positive numbers; however, it is clear that at this stage of the process, as a testing platform, RTF-EXPAR offers the advantage of both speed and simplicity over other NAATs.



Total Population (P+N) 276	Observed Positive (PP) 91	Observed Negative (PN) 185	Accuracy ((TP+TN)/(P+N)) 94%	
Actual Positive (P) 78	**True Positive** (TP) 76	**False Negative** (FN) 2	True Positive Rate/**Sensitivity** (TP/P) **97%**	False Negative Rate (FN/P) 3%
Actual Negative (N) 198	**False Positive** (FP) 15	**True Negative** (TN) 183	False Positive Rate (FP/N) 8%	True Negative Rate/**Specificity** (TN/N) **92%**
Prevalance (P/(P+N)) 28%	Positive Predictive Value (TP/PP) 84%	False Omission Rate (FN/PN) 1%	False Discovery Rate (FP/PP) 16%	Negative Predictive Value (TN/PN) 99%

***Figure 4.25**: Confusion matrix displaying the results of the Orf1ab and N gene patient sample validation study*

4.6 Conclusions

During the course of this work, we have developed a method of isothermally amplifying RNA without the requirement of a RT step. In comparison to existing NAATs, we have demonstrated that RTF-EXPAR possesses a number of advantages - not only is it much faster than its RT contemporaries, but it also holds a similar sensitivity level when it comes to detecting isolated RNA. However, when it comes to crude samples, RTF-EXPAR proves to be vastly superior to not only LAMP but also to the 'gold standard' of PCR, in both time and in limit of detection. This may prove to be a major step forward in assay development, due not only to the speed and simplicity of the assay but also in the design. LAMP is looking to be the next step in NAAT; however, the complexity of the assay could also be its drawback, RTF-EXPAR may yet have time to usurp PCRs crown.

4.7 Future Work

During the course of this work, we have developed a method of detecting SARS-CoV-2 in a matter of minutes. As well as continuing to optimise the conditions, the readout system can also be changed, for example through the inclusion of a probe/quencher system which could create a multiplex system. As we have seen, our **Template COVID´-COVID´** is sequence specific to **COVID Trigger**. Through the inclusion of alternative templates with a distinct fluorophore/quencher tag, multiple pathogens could be screened simultaneously. Such an approach would be extremely beneficial not only for respiratory infections, but also for a plethora of other diseases, such as HIV and cancers. The inclusion of RTF-EXPAR into the M13:DNA system we discussed in Chapter 3 could be an ideal method of detecting various sequences.

The speed of this assay offers a significant advantage over other tests of a similar ilk. HPV is an extremely common virus that can lead to cancer. Preliminary work with Dr Elaine Leung at the Institute of Cancer and Genomic Sciences has shown that RTF-EXPAR was successfully able to detect the presence of HPV RNA - the deployment of a rapid test for HPV in developing countries would have unfathomable benefits to many young lives.

Finally, further optimisation could enable a colourimetric readout - work by NEB[51] has established that the removal of the buffer and inclusion of a pH dye can lead to a change in colour upon amplification. The combination of these abilities could lead to an assay that is not only accurate, but also easy to use that can be deployed in both domestic settings and the developing world, with a simple readout, thus helping to diagnose life-changing afflictions in a matter of minutes.

4.8 References

1. WHO | Pneumonia of unknown cause – China. *WHO* http://www.who.int/csr/don/05-january-2020-pneumonia-of-unkown-cause-china/en/.
2. WHO | Novel Coronavirus – China. *WHO* http://www.who.int/csr/don/12-january-2020-novel-coronavirus-china/en/.
3. Past Pandemics | Pandemic Influenza (Flu) | CDC. https://www.cdc.gov/flu/pandemic-resources/basics/past-pandemics.html (2019).
4. *Molecular cell biology*. (W.H. Freeman, 2000).
5. Payne, S. Family Coronaviridae. in *Viruses* 149–158 (Elsevier, 2017). doi:10.1016/B978-0-12-803109-4.00017-9.
6. Coronavirus | Human Coronavirus Types | CDC. https://www.cdc.gov/coronavirus/types.html (2020).
7. Yip, C. C.-Y. *et al.* Development of a Novel, Genome Subtraction-Derived, SARS-CoV-2-Specific COVID-19-nsp2 Real-Time RT-PCR Assay and Its Evaluation Using Clinical Specimens. *IJMS* **21**, 2574 (2020).
8. Chen, Y. *et al.* A highly specific rapid antigen detection assay for on-site diagnosis of MERS. *Journal of Infection* **73**, 82–84 (2016).
9. Bhadra, S. *et al.* Real-Time Sequence-Validated Loop-Mediated Isothermal Amplification Assays for Detection of Middle East Respiratory Syndrome Coronavirus (MERS-CoV). *PLoS ONE* **10**, e0123126 (2015).
10. Pallesen, J. *et al.* Immunogenicity and structures of a rationally designed prefusion MERS-CoV spike antigen. *Proc Natl Acad Sci USA* **114**, E7348–E7357 (2017).
11. Parks, J. M. & Smith, J. C. How to Discover Antiviral Drugs Quickly. *N Engl J Med* **382**, 2261–2264 (2020).

12. Cascella, M., Rajnik, M., Cuomo, A., Dulebohn, S. C. & Di Napoli, R. Features, Evaluation and Treatment Coronavirus (COVID-19). in *StatPearls* (StatPearls Publishing, 2020).
13. Billah, Md. A., Miah, Md. M. & Khan, Md. N. Reproductive number of coronavirus: A systematic review and meta-analysis based on global level evidence. *PLoS ONE* **15**, e0242128 (2020).
14. World Health Organization. *Consensus document on the epidemiology of severe acute respiratory syndrome (SARS)*. https://apps.who.int/iris/handle/10665/70863 (2003).
15. Kucharski, A. J. & Althaus, C. L. The role of superspreading in Middle East respiratory syndrome coronavirus (MERS-CoV) transmission. *Eurosurveillance* **20**, (2015).
16. Wong, Z. S. Y., Bui, C. M., Chughtai, A. A. & Macintyre, C. R. A systematic review of early modelling studies of Ebola virus disease in West Africa. *Epidemiol. Infect.* **145**, 1069–1094 (2017).
17. Kelly, H. *et al.* The Age-Specific Cumulative Incidence of Infection with Pandemic Influenza H1N1 2009 Was Similar in Various Countries Prior to Vaccination. *PLoS ONE* **6**, e21828 (2011).
18. Fraser, C. *et al.* Pandemic Potential of a Strain of Influenza A (H1N1): Early Findings. *Science* **324**, 1557–1561 (2009).
19. Petersen, E. *et al.* Comparing SARS-CoV-2 with SARS-CoV and influenza pandemics. *The Lancet Infectious Diseases* **20**, e238–e244 (2020).
20. Sichone, J., Simuunza, M. C., Hang'ombe, B. M. & Kikonko, M. Estimating the basic reproduction number for the 2015 bubonic plague outbreak in Nyimba district of Eastern Zambia. *PLoS Negl Trop Dis* **14**, e0008811 (2020).
21. Gani, R. & Leach, S. Transmission potential of smallpox in contemporary populations. *Nature* **414**, 748–751 (2001).
22. Guerra, F. M. *et al.* The basic reproduction number (R 0) of measles: a systematic review. *The Lancet Infectious Diseases* **17**, e420–e428 (2017).

23. WHO Director-General’s opening remarks at the media briefing on COVID-19 - 16 March 2020. https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---16-march-2020.
24. Regional offices. https://www.who.int/about/who-we-are/regional-offices.
25. WHO Coronavirus Disease (COVID-19) Dashboard. https://covid19.who.int/.
26. Malik, Y. S. *et al.* Emerging novel coronavirus (2019-nCoV)—current scenario, evolutionary perspective based on genome analysis and recent developments. *Veterinary Quarterly* **40**, 68–76 (2020).
27. Chen, Z. *et al.* Rapid and Sensitive Detection of anti-SARS-CoV-2 IgG, Using Lanthanide-Doped Nanoparticles-Based Lateral Flow Immunoassay. *Anal. Chem.* **92**, 7226–7231 (2020).
28. Wu, A. *et al.* Genome Composition and Divergence of the Novel Coronavirus (2019-nCoV) Originating in China. *Cell Host & Microbe* **27**, 325–328 (2020).
29. Li, Y., Bai, W. & Hashikawa, T. The neuroinvasive potential of SARS-CoV2 may play a role in the respiratory failure of COVID-19 patients. *J Med Virol* **92**, 552–555 (2020).
30. Huang, W. E. *et al.* RT-LAMP for rapid diagnosis of coronavirus SARS-CoV-2. *Microb. Biotechnol.* **13**, 950–961 (2020).
31. Nguyen, T., Duong Bang, D. & Wolff, A. 2019 Novel Coronavirus Disease (COVID-19): Paving the Road for Rapid Detection and Point-of-Care Diagnostics. *Micromachines* **11**, 306 (2020).
32. Zhu, F.-C. *et al.* Safety, tolerability, and immunogenicity of a recombinant adenovirus type-5 vectored COVID-19 vaccine: a dose-escalation, open-label, non-randomised, first-in-human trial. *The Lancet* **395**, 1845–1854 (2020).
33. Alkan, F. *et al.* The Detection and Genetic Characterization Based on the S1 Gene Region of BCoVs from Respiratory and Enteric Infections in Turkey: Detection and Genetic Characterization Based on the S1 Gene. *Transboundary and Emerging Diseases* **58**, 179–185 (2011).

34. Pradhan, S. K. *et al.* Recombinant nucleocapsid protein based single serum dilution ELISA for the detection of antibodies to infectious bronchitis virus in poultry. *Journal of Virological Methods* **209**, 1–6 (2014).
35. Kashir, J. & Yaqinuddin, A. Loop mediated isothermal amplification (LAMP) assays as a rapid diagnostic for COVID-19. *Medical Hypotheses* **141**, 109786 (2020).
36. Koch, T. *et al.* Safety and immunogenicity of a modified vaccinia virus Ankara vector vaccine candidate for Middle East respiratory syndrome: an open-label, phase 1 trial. *The Lancet Infectious Diseases* **20**, 827–838 (2020).
37. Mahase, E. Covid-19: Innova lateral flow test is not fit for “test and release” strategy, say experts. *BMJ* m4469 (2020) doi:10.1136/bmj.m4469.
38. Wise, J. Covid-19: Lateral flow tests miss over half of cases, Liverpool pilot data show. *BMJ* m4848 (2020) doi:10.1136/bmj.m4848.
39. Deeks, J. J. & Raffle, A. E. Lateral flow tests cannot rule out SARS- CoV-2 infection. *BMJ* m4787 (2020) doi:10.1136/bmj.m4787.
40. Rifkind, D. & Freeman, G. L. *The Nobel Prize winning discoveries in infectious diseases*. (Elsevier/Academic, 2005).
41. Adams, G. A beginner’s guide to RT-PCR, qPCR and RT-qPCR. *The Biochemist* **42**, 48–53 (2020).
42. Viasure SARS-COV-2 RT-PCR Kits available exclusively from Pro-Lab Diagnostics in the UK -. *Pro Lab* https://www.pro-lab.co.uk/2020/03/06/covid-19-ce-marked-kits-available-now-from-pro-lab-diagnostics/ (2020).
43. SARS-CoV-2 Rapid Colorimetric LAMP Assay Kit | NEB. https://international.neb.com/products/e2019-sars-cov-2-rapid-colorimetric-lamp-assay-kit#Product%20Information.
44. Overbergh, L. *et al.* The use of real-time reverse transcriptase PCR for the quantification of cytokine gene expression. *J Biomol Tech* **14**, 33–43 (2003).

45. Corman, V. M. *et al.* Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Eurosurveillance* **25**, (2020).
46. Hindson, B. J. *et al.* High-Throughput Droplet Digital PCR System for Absolute Quantitation of DNA Copy Number. *Anal. Chem.* **83**, 8604–8610 (2011).
47. Didenko, V. V. DNA Probes Using Fluorescence Resonance Energy Transfer (FRET): Designs and Applications. *BioTechniques* **31**, 1106–1121 (2001).
48. NHS Test and Trace: how we test your samples. *GOV.UK* https://www.gov.uk/government/publications/nhs-test-and-trace-how-we-test-your-samples/nhs-test-and-trace-how-we-test-your-samples.
49. Esbin, M. N. *et al.* Overcoming the bottleneck to widespread testing: a rapid review of nucleic acid testing approaches for COVID-19 detection. *RNA* **26**, 771–783 (2020).
50. Zhang, X., Lowe, S. B. & Gooding, J. J. Brief review of monitoring methods for loop-mediated isothermal amplification (LAMP). *Biosensors and Bioelectronics* **61**, 491–499 (2014).
51. Tanner, N. A., Zhang, Y. & Evans, T. C. Visual detection of isothermal nucleic acid amplification using pH-sensitive dyes. *BioTechniques* **58**, 59–68 (2015).
52. Murray, I. A., Stickel, S. K. & Roberts, R. J. Sequence-specific cleavage of RNA by Type II restriction enzymes. *Nucleic Acids Research* **38**, 8257–8268 (2010).
53. Carter, J. G. *et al.* Ultrarapid detection of SARS-CoV-2 RNA using a reverse transcription–free exponential amplification reaction, RTF-EXPAR. *Proc Natl Acad Sci USA* **118**, e2100347118 (2021).
54. NEBcloner. BstNI Restriction Enzyme Digestion. http://nebcloner.neb.com/#!/protocol/re/single/BstNI.
55. Qian, J. *et al.* Sequence dependence of isothermal DNA amplification via EXPAR. *Nucleic Acids Research* **40**, e87–e87 (2012).
56. Oligo Analyzer. https://eu.idtdna.com/calc/analyzer.

57. Peto, T. *et al.* COVID-19: Rapid antigen detection for SARS-CoV-2 by lateral flow assay: A national systematic evaluation of sensitivity and specificity for mass-testing. *EClinicalMedicine* **36**, 100924 (2021).
58. NY 14202274-5487, C. H. 878 M. S. B. RP2 Controls | Shop Respiratory Panel 2 Controls for Assay Developers and Clinical Researchers - ZeptoMetrix Infectious Disease Diagnostics. https://www.zeptometrix.com/products/respiratory-panel-2-rp2-controls-ea.
59. Pouwels, K. B. *et al.* Community prevalence of SARS-CoV-2 in England from April to November, 2020: results from the ONS Coronavirus Infection Survey. *The Lancet Public Health* **6**, e30–e38 (2021).
60. Target Product Profile: Point of Care SARS-CoV-2 detection tests. *GOV.UK* https://www.gov.uk/government/publications/how-tests-and-testing-kits-for-coronavirus-covid-19-work/target-product-profile-point-of-care-sars-cov-2-detection-tests.

Chapter 5: Linear activity of Poly(*p*-aryltriazole) Foldamers

5.1 Introduction

M13 bacteriophage possesses a high aspect ratio and subsequently a large LD signal in flow. This facet, in conjunction with its stability in comparison to other protein-based assemblies (bacteriophages have evolved to live on surfaces and are resistant to desiccation and extremes of temperatures), make it an ideal scaffold for linear dichroism-based biosensors.[1–5] Previous work in the Dafforn group led to the development of a bacterial sensor through the fusion of the g3p protein of the bacteriophage with an antibody binding domain (FB).[1] The presence of FB enables the binding of an IgG moiety, which is able to bind specifically to a protein on the coat of *E.coli* O157. This binding results in a misalignment when subjected to shear flow, therefore causing a decrease in LD signal. However, while M13 displays high functionality, stability and LD signal, production of the M13 itself is the limiting factor. A standard propagation produces only 16 mg of M13 for every litre of media used, therefore hindering the ability of upscaling for manufacturing. Foldamers are a large field within chemistry which takes inspiration from biological structures, thus leading to the generation of high aspect ratio polymers which could provide an alternative abiotic scaffold to M13 bacteriophage.

5.2 Foldamers

Nature relies on macromolecular polymers that, through inter- and intramolecular interactions, assemble into unique structures to perform specific roles.[6] There are three major biopolymer backbones; peptides, nucleic acids and polysaccharides. Each biomolecule employs specific monomeric units throughout the polymer chain, (e.g. amino acids or nucleotides) which are able to form precise higher ordered structures *via* non-covalent interactions (hydrogen bonding, ionic bonding, and van der Waals forces).[7–9] Despite each interaction individually being weak, the combined effect of multiple localised non-covalent interactions creates a microenvironment that results in the biomolecules possessing areas of highly ordered structure.[9] However, non-covalent interactions are not the only driving force in higher order biomolecular structure, other factors include the conformational preference of the backbones, polar interactions and solvation effects.[10] The nature of these structures are of particular interest to organic and biochemists who wish to mimic and manipulate naturally occurring biological systems.[11] As a result, the study of higher ordered structures driven by non-covalent interactions in solution led to the development of the foldamer field of chemistry.[8,12] In general terms a foldamer is an synthetic oligomer that possesses the ability to fold into a conformationally ordered state when in solution, as a result of one or a more non-covalent interactions.[7,8] By designing the monomer units with specific interactions in mind, be that non-covalent or external stimuli, an oligomer with predictable conformations can be synthesised.[7,13] The result is a tuneable molecule that can be employed in further areas of interest such as molecular recognition,[14–16] artificial ion channels,[14,17] virus mimics,[18] and asymmetric catalysis.[19,20] Most foldamers fall into one of two categories, those with an aliphatic or aromatic backbone. Aliphatic groups are typically based upon the amino acid backbone of proteins. However this category is not limited to proteins alone, with research groups developing foldamers based upon oligoureas,[41] azapeptides,[21] carbohydrates,[22] and methyl methacrylate[23] systems.

The other category, aromatic foldamers, is based upon backbones containing, as the name suggests, an aromatic ring. Unlike aliphatic backbones there is no natural homologue for this class; however, they offer the advantage of increased functionality.[13] Within the aromatic group, there are two further sub-classes of foldamer, those whose assembly is driven by hydrogen bonding and those whose folding is driven by solvent effects.[13]

Hamuro *et al* published one of the first known examples of an aromatic foldamer that assembled *via* hydrogen bonding in 1994.[24,25] The work detailed the use of anthranilic acid to form amide bonds and therefore create a structure similar to that of a β-amino acid where intramolecular hydrogen bonding causes the structure to adopt a *trans* geometry, and as a result folding into α helices.[24,25] The field grew from here with multiple backbone scaffolds being investigated. Benzoylureas are a class of foldamers that form hierarchical structure *via* intramolecular hydrogen-bonding and are composed of repeating benzene-acylurea units. This system leads to the creation of helical structures with a large diameter of ~37 Å.[26,27] Zhang *et al* found through the incorporation of ester carbonyls onto the benzene ring *ortho* to the urea group, folding was promoted in chloroform.[28] Meanwhile, it was discovered by Clayden *et al* that the helical conformation adopted was a result of stabilisation as caused by π-π stacking.[29,30]

The second subclass of aromatic foldamer are those which fold based upon their solvent environment. These foldamers are usually constructed of a hydrophobic aromatic backbone with hydrophilic side chains.[12] This structure, unlike the hydrogen bonding counterparts, enables polar solvents to be employed, resulting in solvophobicity as the driving force of folding. As a result of this, aromatic groups will favour a geometry that limits the amount of surface area in contact with the solvent, while promoting π-π stacking thus increasing stability.[10,31,32] Ben *et al* developed a poly(*m*-phenylene) molecule with a chiral polyethylene glycol (PEG) chain in the 5-position. They demonstrated that as the solvent environment transitioned from chloroform to the more polar water, folding was observed.

Through the use of PEG chains they were able to monitor the folding using circular dichroism (CD) spectroscopy.[33] Furthermore, Meudtner reported foldamers composed of a pyridine bisacetylene and a phenyl bisazide with PEG side chains that employed click chemistry forming triazole rings to assemble.[34] Using CD and fluorescence spectroscopy they were able to monitor the change in conformation and determine the conformation changes were driven by steric constrains and π-π stacking between heteroaromatics rings. Interestingly, they also discovered the helicity was pH and anion dependent, with the conformation changing upon addition of acid.[34]

5.2.1 Determination of folding

Regardless of class or subclass, one of the key elements in the development of synthetic foldamers is to gain an understanding of the folding process and the higher order structure that is formed. This can be studied using a range of analytical techniques depending on the state of the foldamer in question. Solid state foldamers such as those detailed by Peters *et al*,[35]and Lister *et* al[36] have been studied using X-ray crystallography.[25,37] This technique enables for the study of the structure of the foldamer in different conformational states, thus determining the definitive structure of the molecule; however, if the foldamer cannot be crystalised, this technique cannot be employed. On the other hand, solution state offers a wide variety of techniques to study the conformational changes as the unfolded and folded oligomers give different spectroscopic signals. A summary of some methods used to characterise folding is outlined below.

Proton NMR spectroscopy (^{1}H NMR) works particularly well for aromatic foldamers as demonstrated by Meunier *et al*[38] and Le Bailly *et al*.[10] While the presence of a polymer system leads to peak broadening[39] due to the presence of repeating units, the peaks will appear in approximately the same place as a monomer unit.[39] By monitoring the aromatic protons (6.5 – 8.5 ppm), it is possible to monitor helical folding. During helical folding, the stacking of the π systems generates increased shielding of the aromatic protons; consequently the signals generated by the aromatic protons move to a lower ppm as they are shifted upfield.[40,41] Further upfield shifting of the aromatic protons, coupled with signal broadening, is indicative of helical foldamers aggregating into columns.[42]

Another useful technique to characterise foldamer formation is UV-vis spectroscopy.[42] as proven by Steinwand *et al*[43] and Lu *et al*.[44] A frequently seen effect, indicative of folding is hypochromicity, defined as a decrease in absorbance at the λ_{max}.[45] This is usually indicative of π stacking, creating an effect homologous to that seen in nucleotides when bases stack during duplex formation.[45,46] This coupled with another common effect, a bathochromic shift of the spectra,[44,47] where the absorbance of the λ_{max} shifts to a longer wavelength, is indicative of a change in conformation.[48] By monitoring the change in spectra brought on by these two spectral events it is possible to track the folding process.

The final technique to be discussed in this section is CD spectroscopy.[49,50] CD reports on local chiral structures that cause local coupling of (electric and magnetic) transition dipole moments. Some foldamers possess chiral centres, which favour one handedness of the helical conformation over the other, this assembly can, as a result, be tracked *via* CD as shown by Cai *et al*[51] and Pfukwa *et al.*[18] If one handedness is favoured during the assembly process a Cotton effect is observed in the CD spectrum in the chromophore region of the foldamer backbone.[18,49,52,53]

5.2.2 Poly (*p*-aryltriazole) foldamers

Our collaborators at Stellenbosch University, the Klumperman group, have produced a series of polymers that are frequently employed by the Dafforn group for membrane protein isolation. Some of these polymers have included Poly(styrene-*co*-maleic acid) (SMA), poly(styrene-*co*-(*N*-(3-*N*′,*N*′-dimethylaminopropyl)maleimide)) (SMI) and poly(diisobutylene-*alt*-maleic acid) (DIBMA).[54,55] These polymers have enabled the efficient extraction of membrane proteins, such as G-coupled protein receptors (GCPR).[54,56,57] One such polymer synthesised by the Klumperman group however, did not find a use for protein extraction. The polymer in question was synthesised by Pfukwa *et al,*[18] and modelled after the Tobacco Mosaic Virus (TMV), (a similar morphology to M13) with a structure based on a *para*-linked aryltriazole system (**Figure 5.1**). Through the inclusion of ethylene glycol side chains on the aryl group, an aromatic foldamer which folds based upon its solvent environment was created.[18]

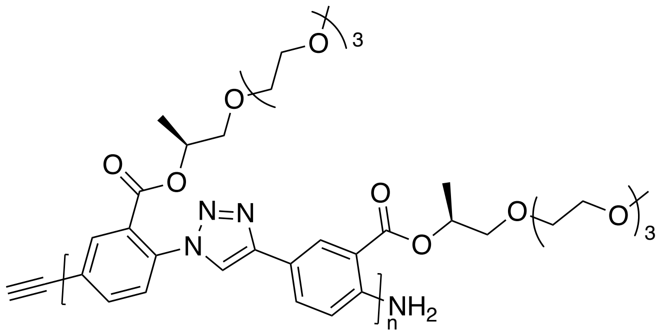


Figure 5.1: *Structure of P(p-AT)* ***S-1***

The seminal work monitored the change in conformation as the solvent system transitioned from organic (DMF) to aqueous (water). As the solvent became increasingly more aqueous, the Poly(*para*-Aryltriazole) [P(*p*-AT)] system folded in a highly ordered assembly process that was monitored using CD and UV-Vis spectroscopy.[18] Briefly explained, at a low water content of 0 – 9% the polymer was fully dissolved and in random orientation, with no CD signal. From 10 – 20% water, the polymer adopts a loosely wound helix, with a positive monosignate CD spectrum. Above 20% water, the polymer adopts a tightly wound helix and stacking of those helices into long fibrous structures take place. The tightly wound helix shows a fully developed bisignate Cotton effect.[18]

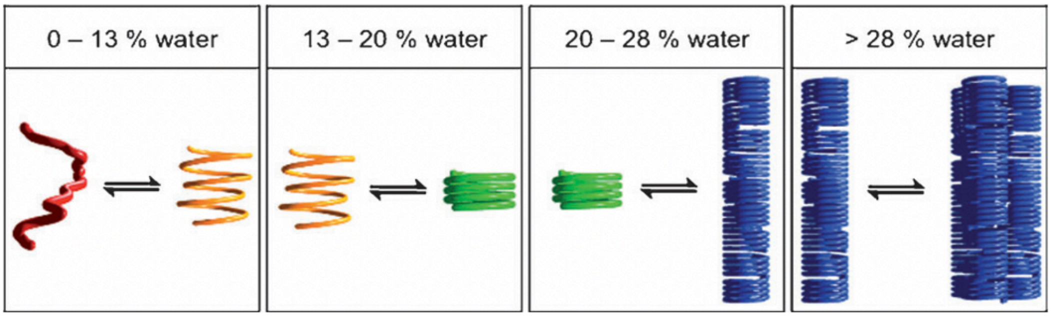


Figure 5.2: *A diagrammatic representation of how the foldamer is expected to assemble in varying solvent conditions Image taken from Pfukwa et al.*[18]

Furthermore, by using molecular modelling simulations, P(*p*-AT) was determined to possess a helical conformation with 14.5 units per turn, resulting in a large diameter of approximately 30.6 Å (compared with M13); this was a direct result of exclusively *para* linked units, thus increasing the rigidity of the backbone and, subsequently, requiring more units for a complete turn.[18] The aryl ring in each unit was predicted to have a triazole ring stacked both above and below, separated by 3.8 Å. The stacking of the foldamers seen at over 28% water, therefore implies that the P(*p*-AT) foldamer possesses a high aspect ratio homologous to biological macromolecules such as DNA and M13 Bacteriophage, which makes the structure amenable to study by LD.[18,58–60] Transmission electron microscopy (TEM) images (**Figure 5.3**) of M13, TMV and P(*p*-AT) display similar structures, further suggesting a high aspect ratio.

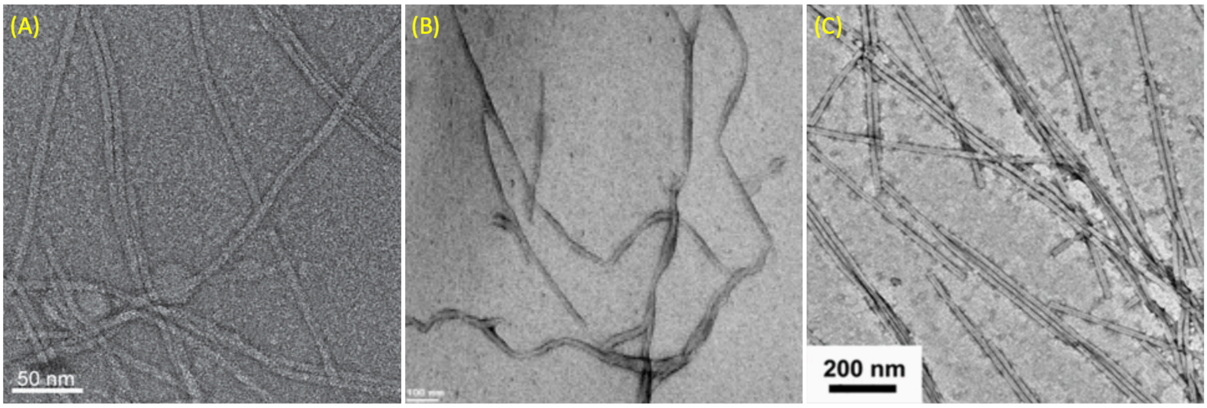


Figure 5.3: *TEM images of (A) M13 Bacteriophage,[61] (B) P(p-AT) Foldamers[18] and (C) Tobacco Mosaic Virus.[62] Images adapted from Niu et al,[61] Pfukwa et al[18] and Liu et al.[62]*

5.3 Aims

We have previously demonstrated that molecules with high aspect ratios, such as the M13 bacteriophage, are an ideal scaffold for LD-based biosensing. Through the conjugation of a biorecognition element, be that DNA, an antibody, or an antigen, we are able to monitor the presence of the complementary analyte in solution. However, the bacteriophage can only be produced with between 20 – 60 mgs of M13 per propagation, and this was after optimisation to maximise the yield.[4] In order to upscale the LD scaffold for commercial use, an alternative scaffold that can be mass produced was required. Chemical synthesis can offer the potential to produce gram scale products, but an LD-active structure is required. Our collaborators in Stellenbosch developed P(*p*-AT) foldamers which have demonstrated the ability to assemble in varying solvent conditions, and as they are based off a TMV, they too should possess a high aspect ratio. The Klumperman group has, over the years, developed a range of polymers that have found interesting uses in the Dafforn group for the extraction of membrane proteins directly from native membranes.[63] However, after developing P(*p*-AT) foldamers in 2013, they are yet to find an application. Therefore, this project sets out to investigate the following:

1) Have the Poly(*para*-Aryltriazole) foldamers synthesised in 2013 degraded?
2) Do the foldamers possess linear activity when analysed *via* LD, and how does that relate to that of M13 bacteriophage?
3) Can LD spectroscopy be used to increase the understanding of foldamer assembly?

5.4 Results and Discussion

5.4.1 Initial Sample Analysis

We were sent a Poly (*para*-Aryltriazole) (P(*p*-AT)) foldamer with a chiral side chain designed and synthesised by our collaborators in Stellenbosch (**Figure 5.1**). Upon receipt of this molecule, the first task was to ensure that structure had not degraded during transit. The most likely degradation site is hydrolysis of the ester bond linking the side chain to the backbone of the polymer, thus resulting in the formation of a carboxylic acid and alcohol as degradation products. As a result, IR was performed to observe which functional groups were present. Both carboxylic acids and esters give a carbonyl C=O stretch around ~1700 cm^{-1} but the former also give a broad OH stretch between 3500 – 2500 cm^{-1} and the latter also give an O-C stretch around ~1000 cm^{-1}.[64,65] Upon examination of the IR spectra of P(*p*-AT) **S-1** (**Figure 5.4**), an absorption was observed at 1720 cm^{-1} to denote a carbonyl group as well as a large peak at 1088 cm^{-1} indicating an O-C bond which may be a result of an ester or the PEG side chain. The peak at 2873 cm^{-1} can be ascribed largely to C-H stretches as it is not as broad as would be expected for a carboxylic acid OH.

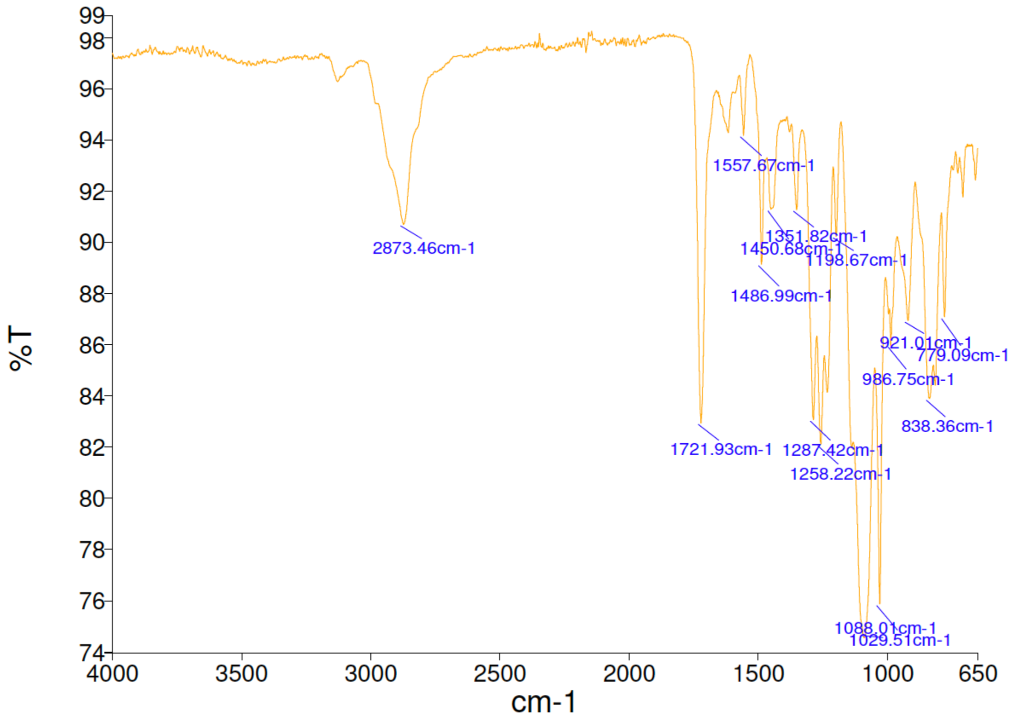


Figure 5.4: *IR spectra of P(p-AT)* ***S-1****, performed on a KBr disc.*

Next a ^{1}H-NMR spectrum was run in d$_6$-DMSO (**Figure 5.5**), which revealed four signals in the aromatic region associated with the aryl (c, d & e) and triazole (b) protons from the foldamer. Furthermore, we were able to see not only the side chains, but also the methyl group (f) adjacent to the ester group responsible for chirality.[18] Therefore both the NMR spectrum and IR spectrum suggested that the foldamer had not degraded in transit.

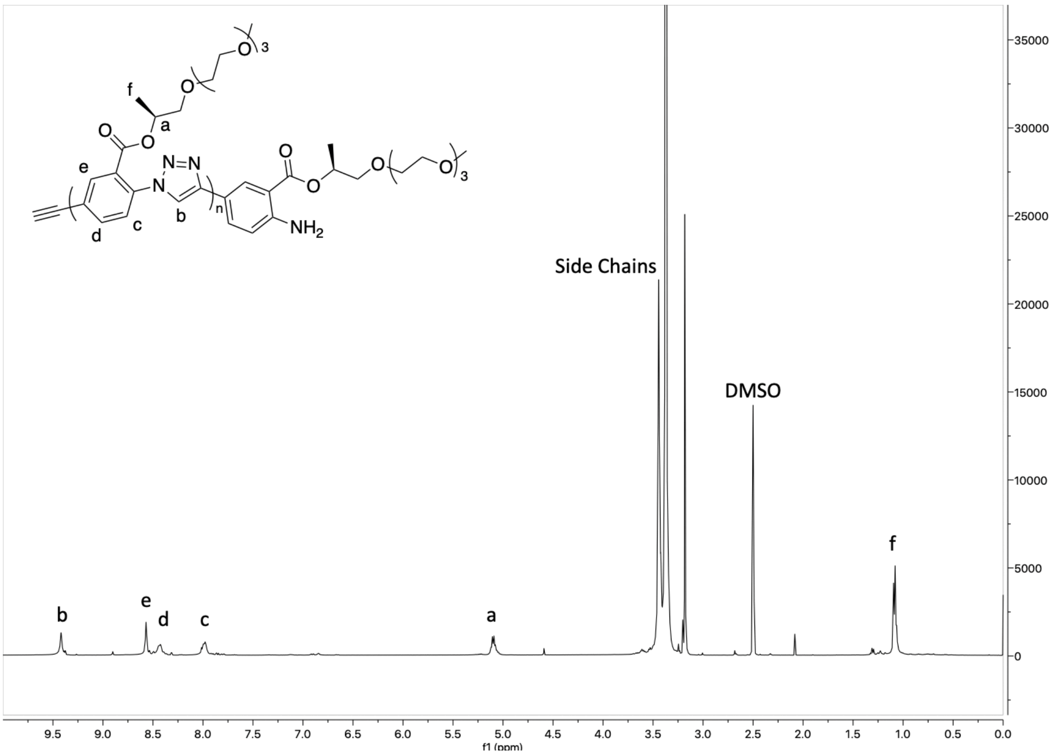


Figure 5.5: *^{1}H NMR spectra of P(p-AT)* ***S-1*** *in DMSO*

As demonstrated using UV-Vis and CD spectroscopy and as previously described, the assembly process of the P(*p*-AT) **S-1** foldamer in changing solvent mixtures (DMF to water) was as follows: From 0 – 13% water, a random coil; up to 20% water, 'loose springs'; up to 28% water, 'tight springs'; >28% water, 'tight springs' stacking to form column like structures.[18] In order to check this assembly process behaviour in the Birmingham labs using the same techniques, numerous samples were composed all using 10 mg of P(*p*-AT) **S-1** in 1 mL of solvent, with the solvent systems ranging from 100% DMF: 0% water to 10% DMF: 90% water.

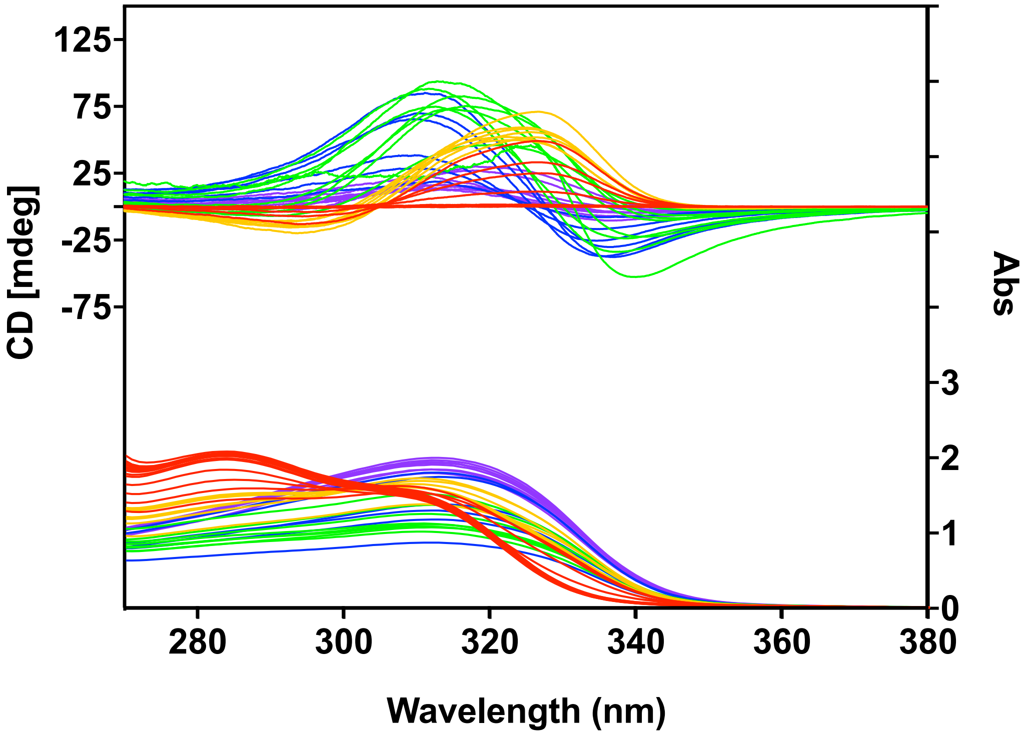


Figure 5.6: *(Top) CD spectra of foldamer P(p-AT) S-1 of concentration 10 μM. (Red) 0 – 13% water, (Orange) 14 – 20% water, (Green) 21 – 28% water, (Blue) 29 – 50% water, (Purple) 55% - 90% water. (Bottom) UV-Vis spectra of foldamer P(p-AT) S-1 of concentration 10 μM. (Red) 0 – 13% water, (Orange) 14 – 20% water, (Green) 21 – 28% water, (Blue) 29 – 90% water.*

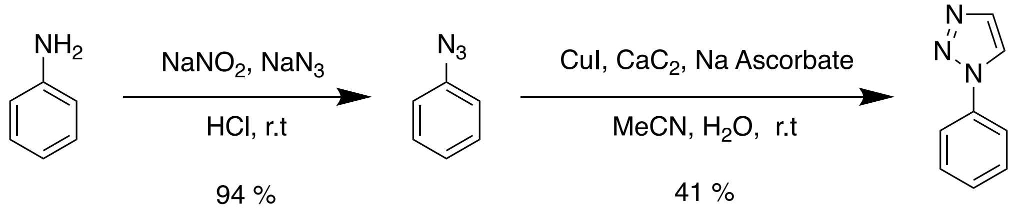
Upon examination of both the UV-Vis and CD spectra (**Figure 5.6**), distinct trends were observed that aligned with the previous work.[18] In the UV-vis, from 0 – 13% water the foldamer possesses a random coil structure which is characterised by a major peak at 284 nm along with a shoulder peak at 310 nm. As the water increases, the peak at 284 nm experiences a hypochromic effect, while the peak at 310 nm experiences hyperchromism, with the intensities being roughly equal at 13% water, the point in which helices start to form. As the water percentage increases still, the 310 nm transitions to become the major UV-vis peak, while 284 nm decreases to become the secondary shoulder peak. The transition of the λ_{max} to a slightly longer wavelength is consistent with increased π stacking. Meanwhile in the CD spectra, at 0 – 9% water there is no signal, indicative of a random structure with no overall helicity. As the water percentage increases, P(*p*-AT) **S-1** folds to create a helical structure, controlled by the hydrophilic side chains and hydrophobic backbone. At 10% water, a single positive Cotton effect is observed, which increases with intensity as the water increases. As the water continually increases to 20% the single Cotton effect evolves into a bisignate signal with negative exciton chirality indicating that a left-handed helix is adopted. The increase in amplitude of the CD couplet between 20–28% water is explained as a further tightening of the stable helical conformation. As the water increases above 28%, while there is no change to the spectral structure, there is a decrease in magnitude. It was originally suggested that at high water content, the intramolecular structure of the helices stays intact, but that the molecules now start assembling into a tertiary structure, that is, nanotubes, thus reducing the magnitude of the CD signal.[18]

Ultimately, both the CD and the UV-Vis spectra produced results consistent with the previous findings, with P(*p*-AT) **S-1** still able to assemble as previously described. With this information to hand, we were able to proceed to the LD studies.

5.4.2 Determining the molecular alignment of chemical moieties during foldamer assembly

Given that M13 bacteriophage produces a large LD signal,[2,3] a key element of using P(*p*-AT) as a replacement scaffold is that the molecule generates an LD signal that is just as intense. LD signals are directly related to the orientation of electronic transitions in the aligned molecule. This means that if the electronic transition observed by LD can be attributed to a specific moiety within the assembly, then its orientation in the assembly can be determined. Therefore, before we investigated the foldamer as a whole, we first were required to investigate the LD signal generated from a backbone subunit, to determine the longest wavelength transition and subsequently the alignment in the moiety using stretched film LD. As previously discussed, this technique enables small molecules, that do not align in flow, to be analysed using LD by exploiting the structure of polyethylene (PE).[66] PE is microcrystalline and therefore upon stretching becomes an ideal environment for orientating molecules in the same direction, either through the adsorption to polymer chains or by occupying anisotropic cavities based on the properties of the analyte being analysed.[66,67] Using this method the molecules are forced to align along their long axis; therefore all measurements are carried out in the parallel plane, horizontal to the incident light. While the foldamer possesses a side chain, we were interested in only the linear signal generated from the backbone itself. Therefore, it was decided to synthesise two backbone subunits in order to determine their alignment. These would not contain side chains as their presence might affect the alignment on stretched film, thereby hindering the ability of full alignment.

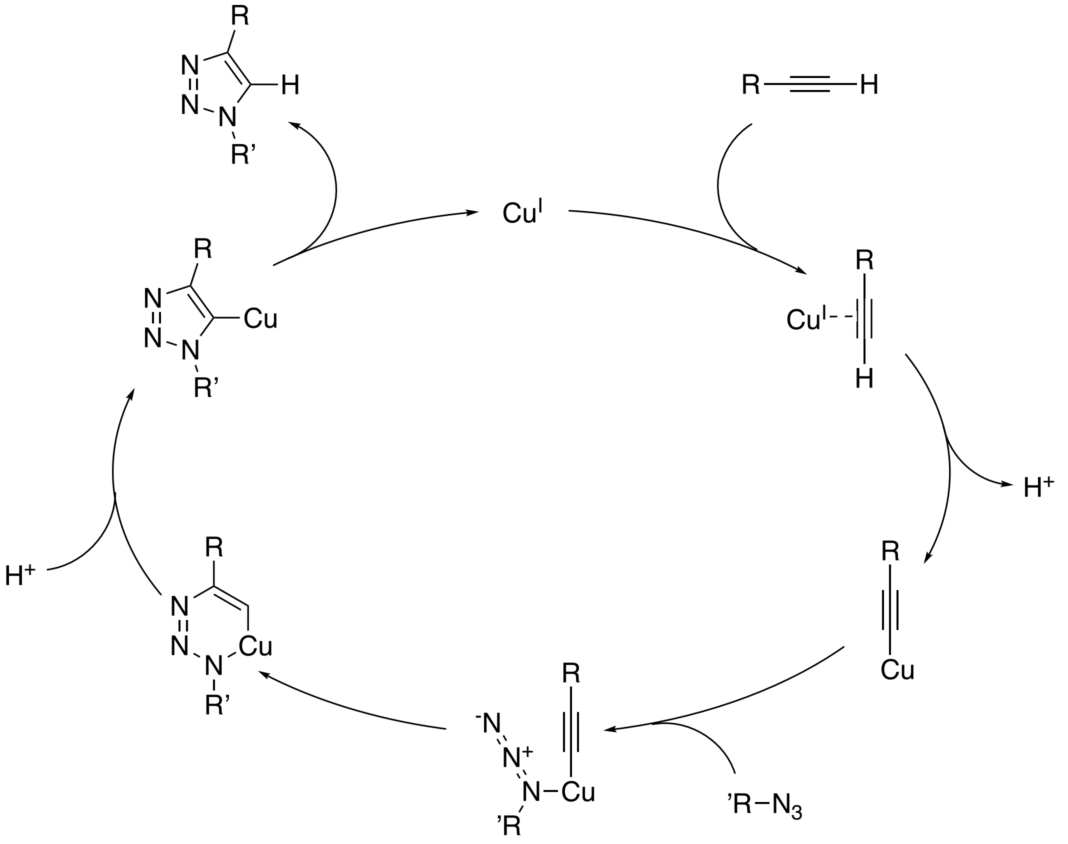
5.4.2.1 Synthesis of single backbone subunits



5.1 **5.2** **PT**

***Scheme 5.1**: Two step synthesis of the single backbone subunit **PT**.*

The single backbone subunit, (**PT**), was prepared (**Scheme 5.1**) using a two-step synthesis beginning with the conversion of aniline (**5.1**) into phenyl azide (**5.2**).[68] Through the reaction of aniline with sodium nitrite in the presence of a strong acid at room temperature a diazonium salt is formed, which is subsequently reacted with sodium azide resulting in the formation of phenyl azide in an excellent yield. The exact mechanism for the conversion of the diazonium salt into the azide is not fully understood.[69] Following the production of phenyl azide, 1-phenyl-1*H*-1,2,3-triazole (**PT**) was produced using a copper catalysed Huisgen cycloaddition reaction (CuAAC).[70]



Scheme 5.2: *Catalytic cycle for the copper catalysed cycloaddition of azide and alkyne.*[71]

An acetylene moiety was generated *in situ* through the reaction of calcium carbide in water. The formation of the salt calcium hydroxide drives the reaction, with the by-product being the release of acetylene, which was subsequently captured by the copper catalysis[72] and fed into the catalytic cycle (**Scheme 5.2**) where (**PT**) was produced in fair yields (41%) following purification. The formation of the triazole ring, and thus the product, was evident from the presence of signals for the aromatic protons at 8.00 ppm and 7.85 ppm (**Figure 5.7**). Further analytical data can be found in Chapter 6.5.6.2.

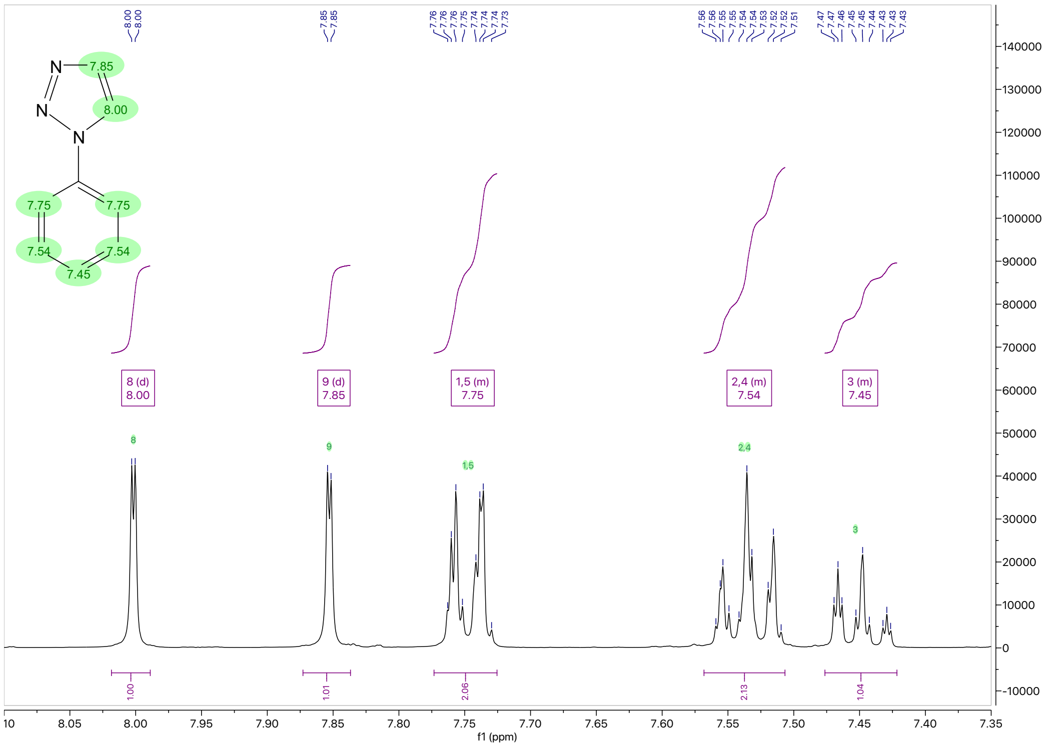
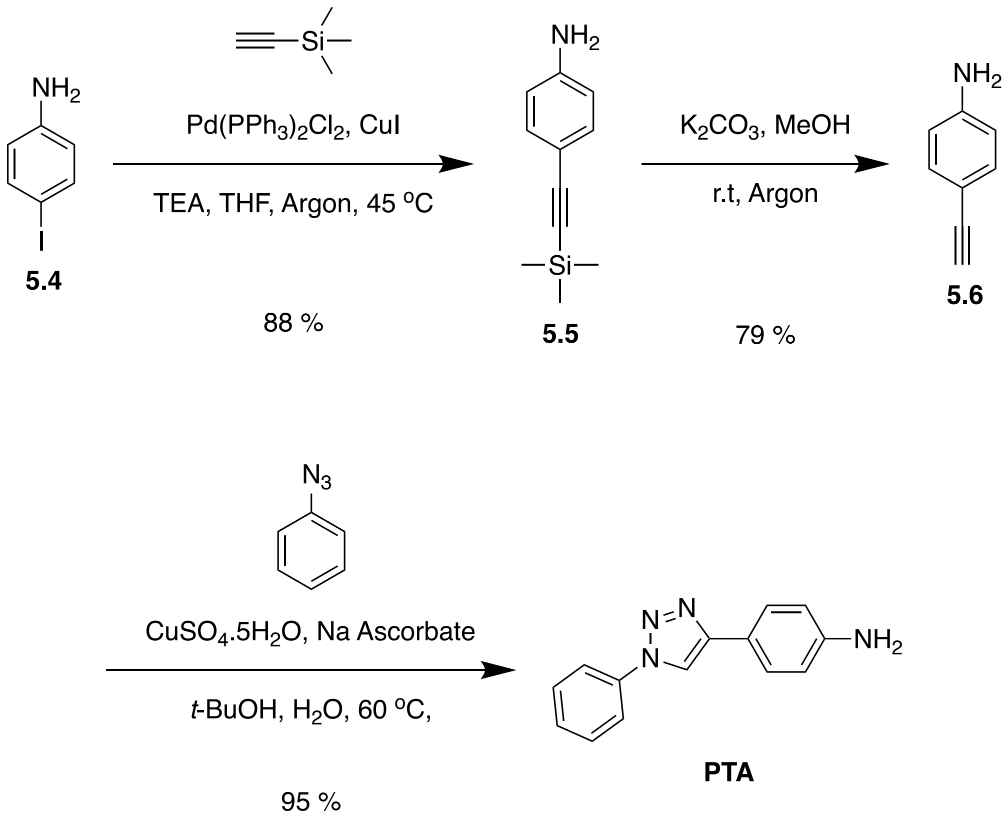


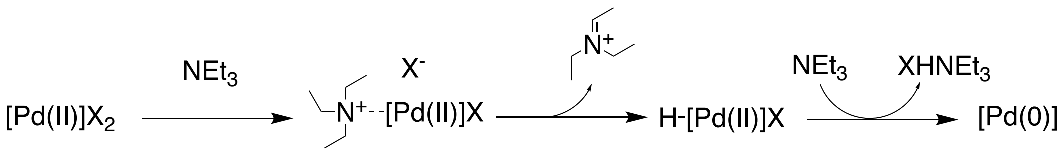
Figure 5.7: *^{1}H NMR spectra of 1-phenyl-1H-1,2,3-triazole* ***(PT)*** *in $CDCl_3$*

5.4.2.2 Synthesis of multiple backbone subunit.



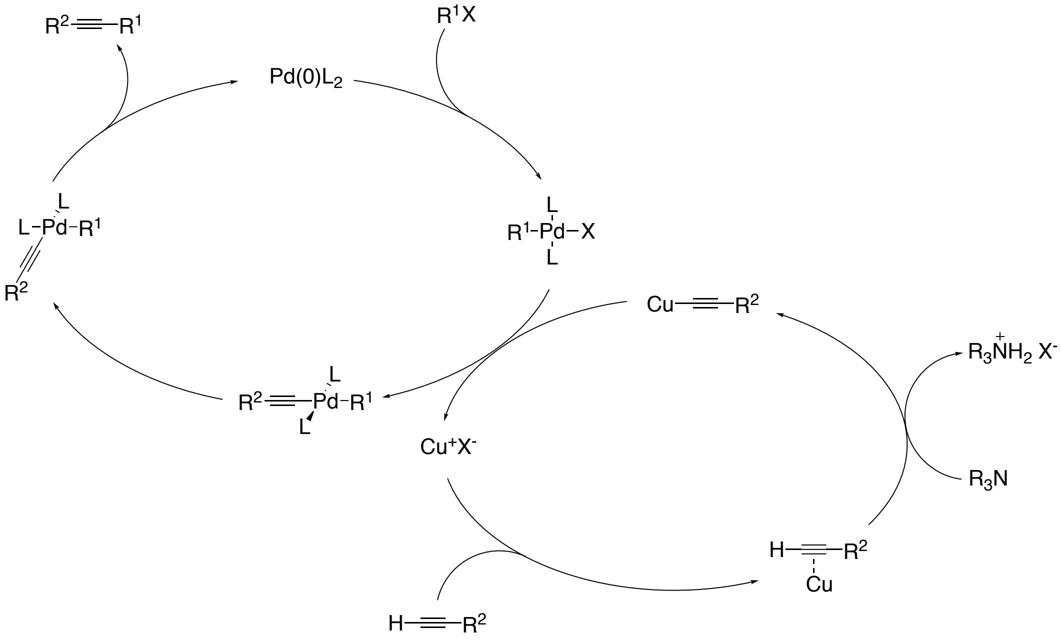
Scheme 5.3: *Three step synthesis of the multiple backbone subunit* ***PTA***

The second backbone subunit, (**PTA**), was synthesised using a three-step procedure. Using a Sonogashira cross coupling reaction, 4-iodoaniline (**5.4**) was converted into 4-((trimethylsilyl)ethynyl)aniline (**5.5**).[73] Sonogashira reactions are duel-catalysed with a copper (I) catalyst and a palladium (0) catalyst. The palladium (II) catalyst Bis(triphenylphosphine)palladium(II) dichloride was used in place of a palladium (0) one due to the instability in air of the latter.[74] Pd(II) is reduced to Pd(0) *in situ* in the presence of triethylamine via a σ-complexation-dehydropalladation-reductive elimination reaction (**Scheme 5.4**).[75]



Scheme 5.4: *σ-complexation-dehydropalladation-reductive elimination of a Pd(II) complex with an amine.*

The reduced palladium complex enables the catalytic cycle to begin (**Scheme 5.5**). The cycle begins with an oxidative addition of the aryl halide with Pd (0), forming a four-coordinated palladium complex. Following the formation of this complex, the palladium cycle intersects with the copper cycle. During the copper cycle a π-alkyne-copper complex is formed in order to increase the acidity of the alkyne, thus aiding deprotonation. Following deprotonation, a copper acetylide is formed, which, interacts with the palladium cycle to form a palladium acetylide via transmetalation. The palladium acetylide undergoes *trans/cis* isomerisation, followed by reductive elimination yielding the crosslinked product and restoring the Pd(0) catalyst.



Scheme 5.5: *Catalytic cycle for a Sonogashira Cross Coupling reaction.*

The Sonogashira reaction yielded 4-((trimethylsilyl)ethynyl)aniline (**5.5**) in very good yields (88%) following purification. In order to expose the alkyne group for cyclisation, a desilylation reaction was performed using potassium carbonate in methanol. The desilylation reaction yielded 4-ethynylaniline **(5.6)** in good yields (79%). As with the synthesis of (**PT**), in order to produce the desired compound **PTA**, we performed a CuAAC reaction between phenyl azide (**5.2**) and 4-ethynylaniline (**5.6**).[70,76] The reaction yielded 4-(1-phenyl-1*H*-1,2,3-triazol-4-yl)aniline (**PTA**) in excellent yield (95%) without the requirement for purification. As with **PT**, we were able to determine successful formation of the desired target through examination of the ^{1}H NMR spectrum (**Figure 5.8**), which gave a signal at 8.99 ppm corresponding to the single proton on the triazole ring. Further analytical data can be found in Chapter 6.5.6.5

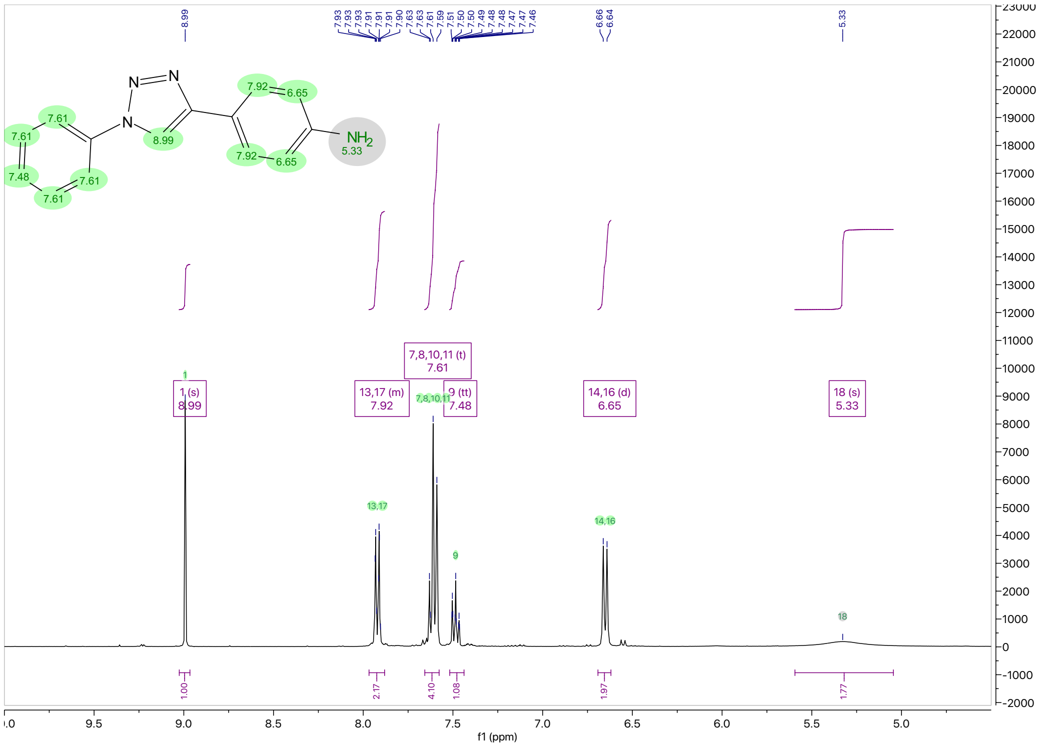


__Figure 5.8__: ^{1}H NMR spectra of 4-(1-phenyl-1H-1,2,3-triazol-4-yl)aniline __(PTA)__ in DMSO

5.4.3 Stretched film LD of backbone subunits.

Upon successful synthesis of the two monomers, **PT** and **PTA**, stretched film LD was then used to determine their **μ**.[67] While the use of PE films for the orientation of molecules has been employed for decades, the hydrophobic nature of the films does limit their ability to align polar molecules. Therefore, by using a technique developed by Razmkhah *et al*, we were able to oxidise the surface of the PE film (Chapter 6.5.5.2), thus creating a hydrophilic environment for the adsorption of molecules.[66] Using PE^{OX} film we were able to characterise the **μ** of the monomers and subsequently use this data to deduce the orientation of polymers in flow LD.[77] Using a mechanical film stretcher (Chapter 6.5.5.2.1), the PE^{OX} films (Chapter 2.2.3.2.2) were stretched prior to the adsorption of molecules. Stretched PE^{OX} forces molecules to orient approximately along the long axis of the molecule. As we are able to orientate the film stretcher, so the long axis is parallel to the laboratory horizontal axis, a positive LD signal indicates that polarisation of the transition is less than 54.7° from this axis. Furthermore, using the results of the stretched film LD studies we are able to assign positive LD to *z* polarisations.[66,67,78]

In order to analyse *via* stretched films, the analyte must be adsorbed onto the surface of the film; this can be achieved in a number of ways; in this case, the simplest method, evaporation, was used.[66,67,78] **PT** was dissolved in chloroform and dropped onto the surface of the film, where the chloroform was allowed to evaporate. This process results in the deposition of monomers on the film, which are forced into alignment with the grooves created by the microcrystalline structures. Unfortunately, **PTA** has limited solubility and was only sufficiently soluble in DMSO, a solvent not known for its volatility. However, as the sample size was less than 100 µL, the evaporation technique was still performed, and while the DMSO took longer to evaporate, the same effect was observed. Following adsorption of the sample onto the film, the LD spectra was recorded, as was the absorbance spectra of film; this was performed using a standard unpolarised spectrophotometer.

Following the acquisition, we were required to transform the LD spectra into the reduced LD (LD^r) spectra. LD spectroscopy is constrained by the same parameters used in other absorbance techniques, such as concentration, pathlength and the dipole strength of the transition. LD^r on the other hand removes these constraints and is consequently a result of only the arrangement of the **μ** relative to the orientation axis.[67] Conveniently, the calculation to obtain the LD^r is a simple one, as stated in **Equation 5.1**, where A_{iso} is the isotropic absorbance of the molecule of interest.[67]

$$LD^r = \frac{LD}{A_{iso}}$$

***Equation 5.1:** Equation to calculate the Reduced Linear Dichroism at a particular wavelength*

While stretch film LD provides a molecule aligned along a single axis, what it does not provide is a randomised sample on a film; therefore, using the LD and the absorbance (A) recorded using unpolarised light, we can used the **Equation 5.2** to calculate A_{iso}.[67]

$$A_{iso} = A + \frac{LD}{3} + \log(1 + 10^{-LD}) - \log 2$$

***Equation 5.2**: Equation to calculate the isotropic absorbance at a particular wavelength.[67]*

Following on from the calculation of A_{iso} we can use **Equation 5.2** to calculate LD^r which, as evident from **Figure 5.9**, displays a signal with the same profile, albeit with a larger magnitude than the LD alone. From the LD^r we can see that the largest signal is generated at 249 nm, with a value of 0.65, which we used to calculate the *S* value. **Equation 5.3** shows that the LD^r is equal to $\frac{3}{2}S(3cos^2\alpha - 1)$, where α is the angle of the **μ** away from the alignment axis, *z*, and *S* is the orientation factor, where 1 is a perfectly aligned sample and 0 is random orientation.

$$LD^r = \frac{3}{2}S(3cos^2\alpha - 1)$$

Equation 5.3: *Equation to calculate the reduced LD by factoring in the alignment factor and angle of the* ***µ*** *from the alignment axis.*

By using **Equation 5.3** we can calculate *S* to be 0.22. While this initially appears to closer to random orientation, this is a good value in terms of small organic molecules. For example anthracene, a standard for film LD, also possesses an S of ~0.22.[66]

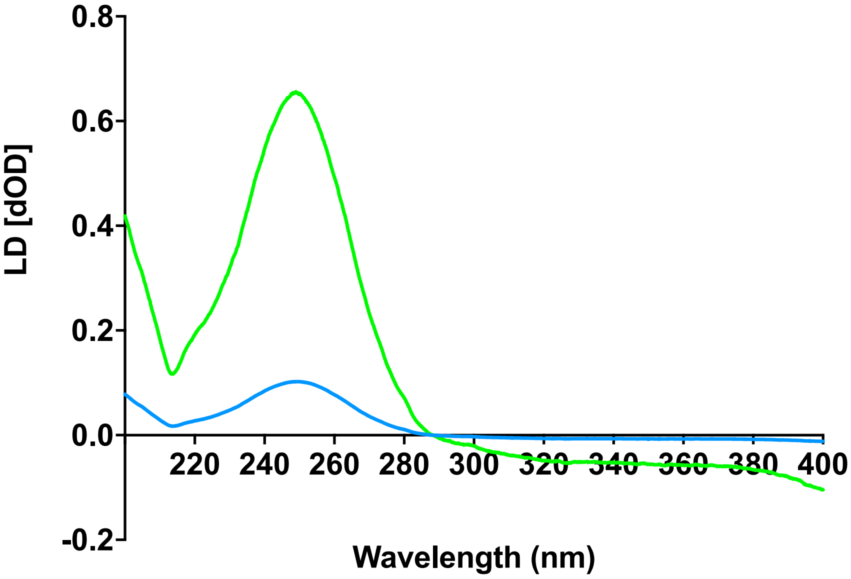


Figure 5.9: *Linear Dichroism spectra generated from the film LD of 1-phenyl-1*H*-1,2,3-triazole (Blue). Reduced Linear Dichroism spectra generated from the film LD of 1-phenyl-1*H*-1,2,3-triazole (Green).*

Using the LD^r and the *S* value we can calculate the angles of other transitions. The transitions are evident by a peak minima or maxima. In **Figure 5.9** there are only the two transitions, the maxima, which we have already discussed and a minima at 213 nm. By using **Equation 5.3** and substituting in *S* we can calculate that α for this transition is 47.8°. Using the two transitions we are able to predict the **µ** angles onto the structure of the molecule.

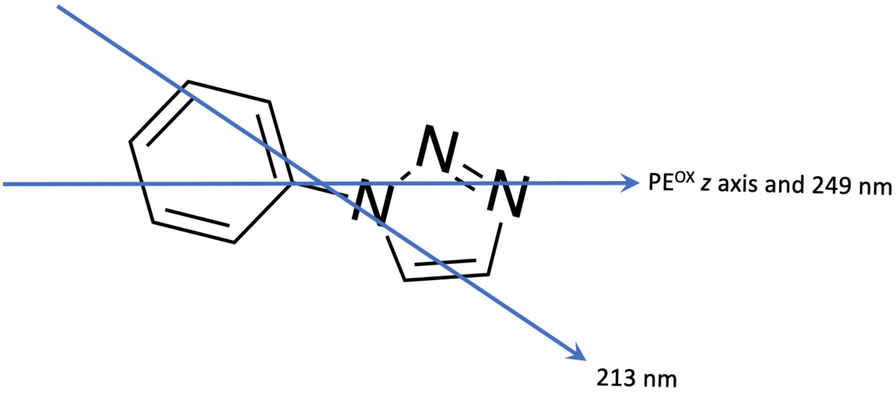


Figure 5.10*: Predicted **μ** of 1-Phenyl-1H-1,2,3-triazole.*

The same process can be repeated for the second of the monomers, **PTA**, this monomer has the maximum LD^r at 287 nm with a value of 0.52 (**Figure 5.11**). Using the same method as **PT** we know that the largest transition is a result of the long axis aligning with the stretch direction (*z*). Therefore using **Equation 5.3** we were able to calculate *S* to be 0.18, which is comparable to the alignment of **PT**.

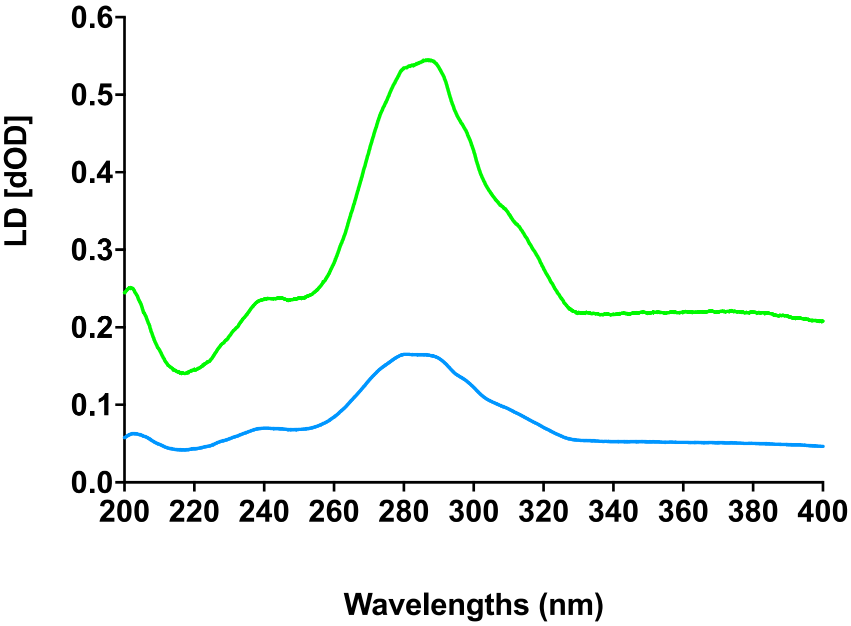


Figure 5.11: *Linear Dichroism spectra generated from the film LD of 4-(1-phenyl-1H-1,2,3-triazol-4-yl)aniline (Blue). Reduced Linear Dichroism spectra generated from the film LD of 4-(1-phenyl-1H-1,2,3-triazol-4-yl)aniline (Green).*

A secondary transition is visible at 218 nm, by inputting the *S* that was previously calculated, we were able to determine the angle of transition to be 44.7° (**Figure 5.12**). 44.7° is only 10° away from the ‘magic angle, which account for LD signal approaching 0 yet still being positive.

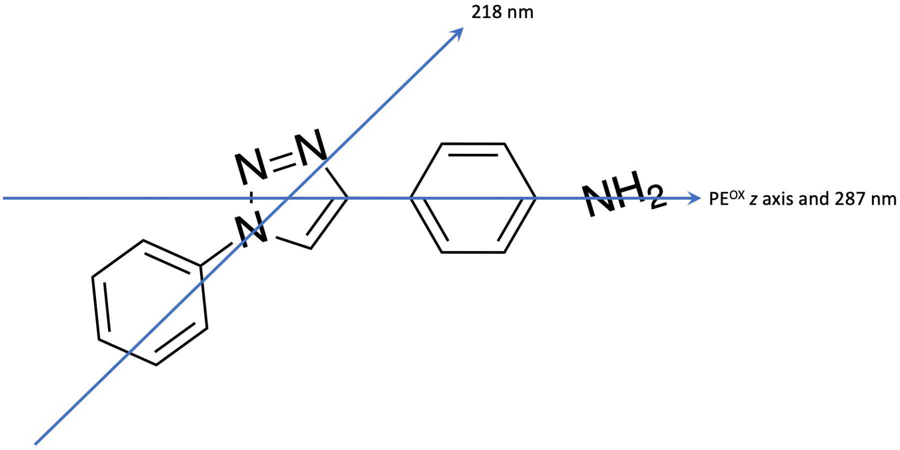


Figure 5.12*: Predicted* **μ** *of 4-(1-phenyl-1H-1,2,3-triazol-4-yl)aniline.*

The results of our film LD studies indicated that the **μ** is along the long axis of the molecule, and not a result of π – π stacking between multiple molecules. This provides us with information that can be transferred for interpretation of the molecular assembly as described in the next section.

5.4.4 Linear Dichroism of Foldamers

As already described above, M13 bacteriophage produces a large LD signal (**Figure 5.13**); therefore, in order to use P(*p*-AT) as a replacement, the signal must also be large.

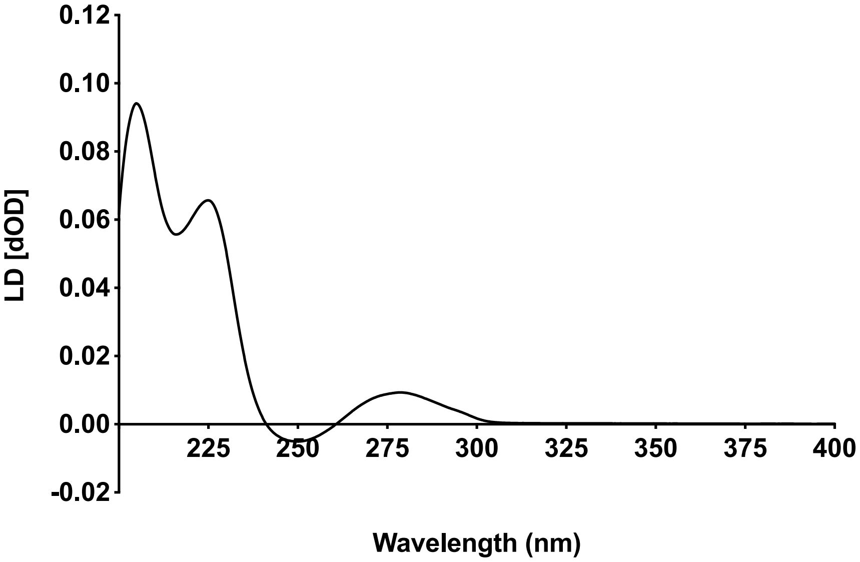


***Figure 5.13**: LD signal generated form M13 bacteriophage in 1x Phosphate Buffered Saline (PBS) pH 7.4 at room temperature.*

With the structure of the foldamer and assembly process analysed and subsequently determined to be conserved alongside the **μ** of the backbone, we were able to begin analysis of the molecule using LD. It is worth noting that chirality has no effect on the LD signal. Using samples of the same composition as during our CD and UV-Vis studies we were able to determine if P(*p*-AT) **S-1** possessed linear properties. As previously described, LD signals are caused by the alignment of materials with respect to one another, therefore, it is expected that although the random coil has an extended conformation, its LD signal in flow would be negligible as the chain lacks order. As the chain assembles into a supramolecular structure, the structure becomes more ordered, and the morphology becomes ever more that of a high aspect ratio structure (e.g., rod-like); as a result, we would expect that these changes in morphology should provide a change in LD signal. Initially the LD spectra (**Figure 5.14**) acted as expected, displaying no signal in the random coil form. The lack of LD signal indicates that there is no molecular organisation when there is less than 10% selective solvent, since the molecularly dissolved polymers do not align in the laminar flow field.

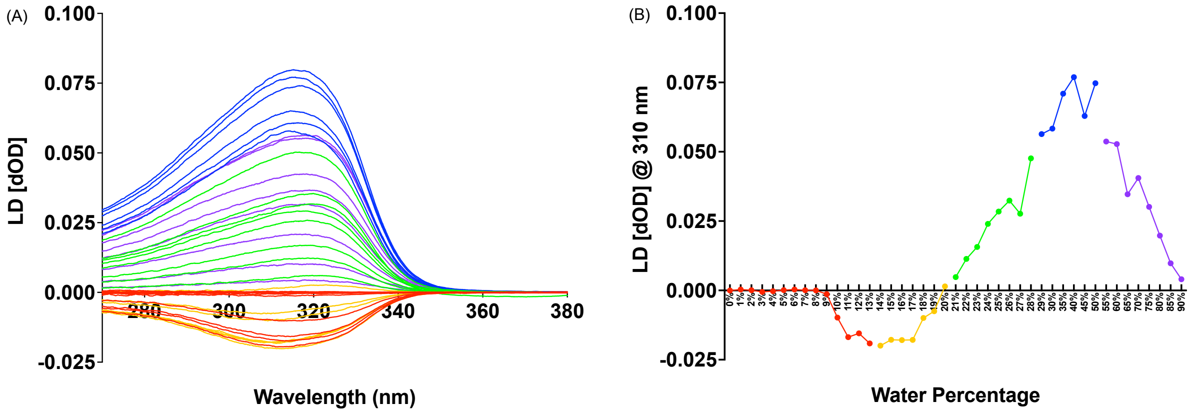


Figure 5.14: *(Left) LD spectra of foldamer P(p-AT) S-1 of concentration 10 µM. (Red) 0 – 13% water, (Orange) 14 – 20% water, (Green) 21 – 28% water, (Blue) 29 – 90% water. (Right) 310 nm wavelength analysis of LD spectra of foldamer P(p-AT) S-1 of concentration 10 µM. (Red) 0 – 13% water, (Orange) 14 – 20% water, (Green) 21 – 28% water, (Blue) 29 – 50% water, (Purple) 55% - 90% water.*

From 10 – 19% water a negative LD signal is observed; based upon the **µ** of the monomers, this indicates that the long axis of the backbone monomer is aligned more perpendicular to the flow axis than the long axis of the polymers overall structure itself. This can be visualised as the azole and phenyl rings stacking in a fashion homologous to DNA bases in dsDNA where the alignment is along the axis of the DNA. Using **Equation 5.3** we are able to determine that the LD^r (**Figure 5.15**) at 330 nm is -0.23, which, using **Equation 5.3**, equates *S* to be 0.07. Flexible polymers such as a strand of duplex DNA 500 nm long has an *S* of ~0.01.[79] This implies that P(*p*-AT) **S-1** is a relatively flexible polymer.

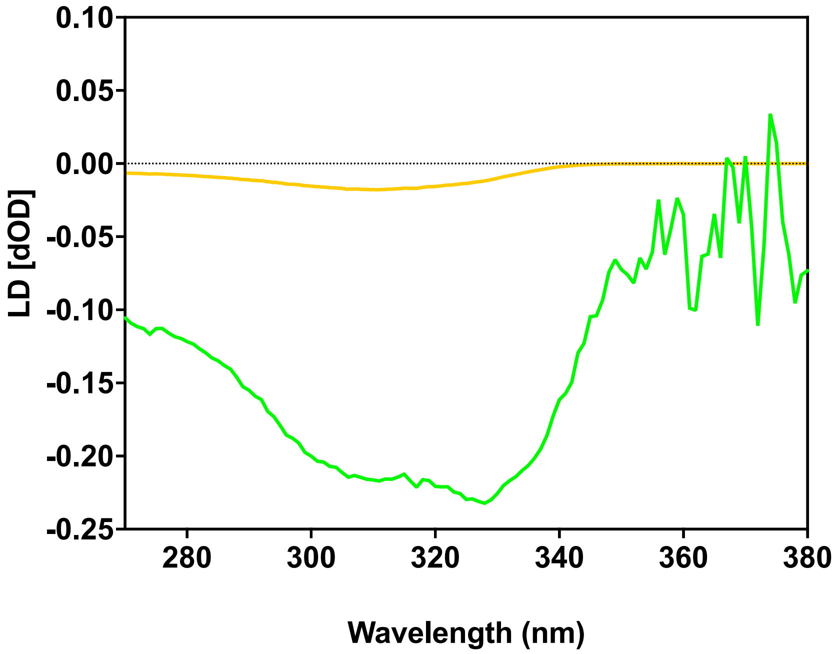


Figure 5.15: *Linear Dichroism spectra generated from the flow LD of P(p-AT)* ***S-1*** *in 17% water (orange). Reduced Linear Dichroism spectra generated form the flow LD of P(p-AT)* ***S-1*** *in 17% water (Green).*

At 20% water, the LD signal returns to zero, suggesting that the foldamer has changed conformation, with the long axis of the monomer transitioning from perpendicular from the alignment axis, to now being equal to an angle of 54.7°, away from the foldamer long axis. According to **Equation 5.3**, this would result in a LD signal equal to zero.[67,80] Above 20%, the LD signal becomes positive and increases in magnitude as the concentration of water increases. Based upon the theory of how the molecule folds, this occurs at the transition from loose coil to tight coil (like the arrangement of the pVIII protein in M13) however, the change in sign of the LD signal is indicative of a different orientation of the aromatic chromophores. The **μ** of the backbone indicates that above 20% water, the long axis of the backbone now aligns parallel to the foldamer long axis. This is unfeasible based upon the original findings,[18,81] suggesting another assembly or conformation that cannot simply be explained by a tightening of a coil.

As the water continually increases to above 45% water, the LD signal begins to decrease for a second time. Pfukwa *et al* previously theorised that as the water concentration further increases, the tight coils of the foldamer stack into columns, then group together, being driven by the hydrophobicity of the backbone of the molecule.[18] However, this once again does not fit with the results of the LD; the decrease in LD signal signifies the **μ** is transitioning away from being parallel to the alignment axis of the foldamer back to being perpendicular. Interestingly, while the signal did decrease in magnitude at higher water percentages, the signal did not return to zero or change sign, indicating that the **μ** does not transition through the 'magic angle'. While the CD and absorbance studies do indicate folding is happening, the exact manner in which P(*p*-AT) folds do not appear to be as previously described. By including the results of our LD studies, we tentatively assign this observation to the formation of elongated superstructures composed of tighter stacked helices that intertwine to form a 'helix of helices', also known as superhelices. McLachlan *et al*[79] demonstrated that a rigid structure (such as M13 bacteriophage) with a length of ~1 μm or longer (as we predict the foldamer to be) possesses a theoretical *S* of ~0.6 when subjected to shear flow. At 40% water (**Figure 5.16**), the LD^r of nm is 325 nm is 1.47; using the theoretical *S* value from McLachlan[79] in **Equation 5.3** we can determine that **μ,** and thus the monomer long axis, is titled 20.7° from the along axis of the helix. These superhelices would explain not only the folding observed in CD and UV-Vis but also the change of the **μ** evident in the LD. That being said, the decrease in LD above 45% water is still perplexing.

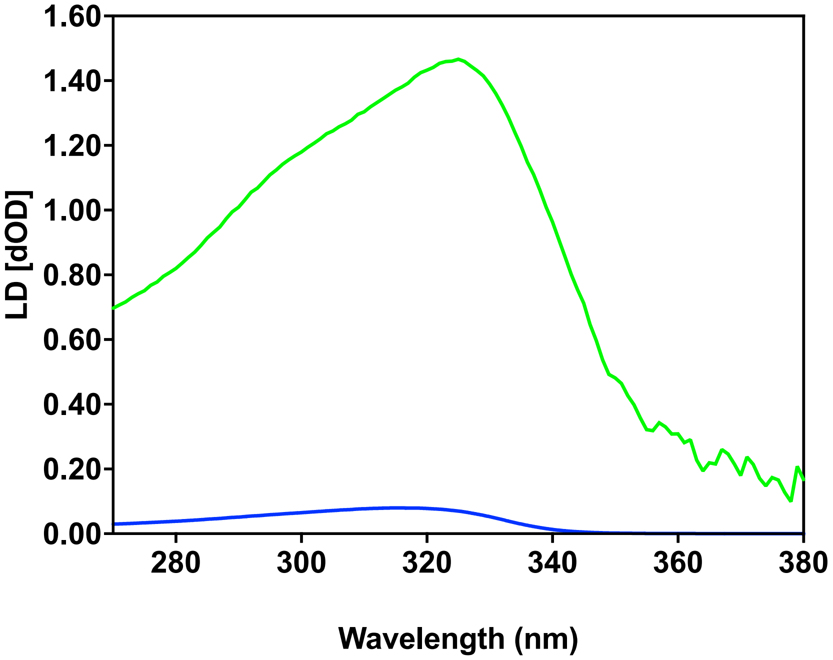


***Figure 5.16**: Linear Dichroism spectra generated from the flow LD of P(p-AT) **S-1** in 40% water (Blue). Reduced Linear Dichroism spectra generated form the flow LD of P(p-AT) **S-1** in 40% water (Green).*

From the LD data we can see that P(*p*-AT) possesses a large LD signal comparable to the signal generated from M13 bacteriophage; however, the changes in LD signals indicate a more complicated change in conformation that previously gleamed from the UV-vis and CD spectra. We therefore endeavoured to probe this structural change further, as described below.[81]

5.4.5 Transmission Electron Microscopy Studies

The findings of our LD studies suggest that the initial assignment *via* CD and UV-Vis was not the complete picture. The change in LD signal implied that the **μ** of the foldamer goes from no overall direction, to perpendicular to the alignment axis, to parallel to the alignment and finally shifting away from the parallel, possibly due to the formation of superhelices. To investigate this further, we reached out to a Professor Corinne Smith at the University of Warwick to perform microscopy studies. Four different solvent domains all of the same concentration (10 mg/mL) were examined (**Figure 5.17**):

1) 0% Water – No LD signal
2) 15% Water – Negative LD Signal
3) 45% Water – Positive LD Signal
4) 90% Water – Decreased Positive LD Signal

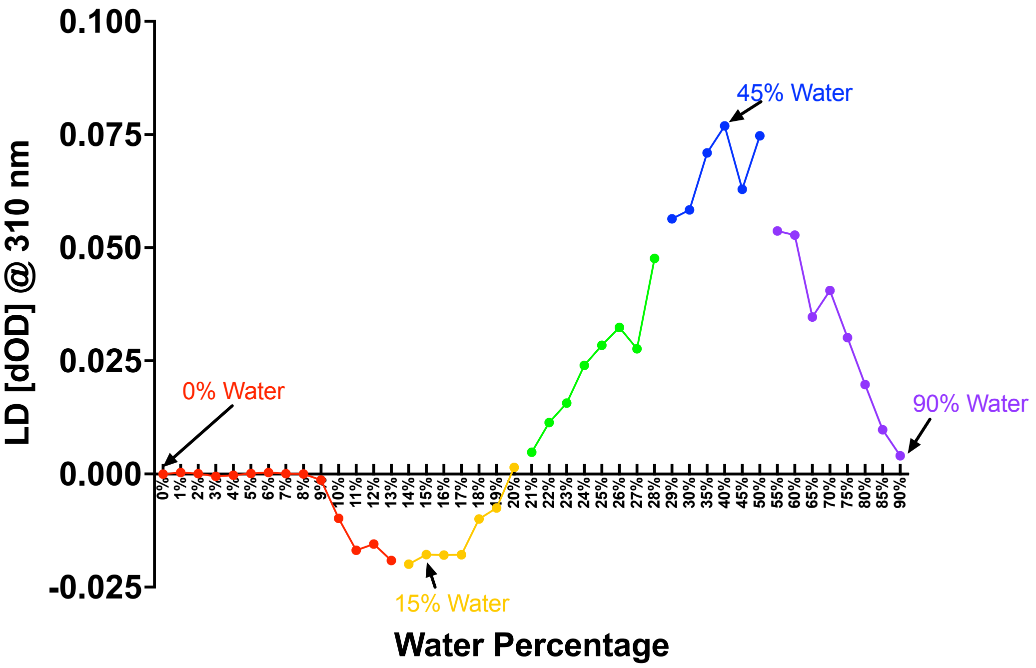


Figure 5.17*: 310 nm wavelength analysis of LD spectra of foldamer P(p-AT) S-1 of concentration 10 μM, with the water percentages used in TEM labelled. (Red) 0 – 13% water, (Orange) 14 – 20% water, (Green) 21 – 28% water, (Blue) 29 – 50% water, (Purple) 55% - 90% water.*

All microscopy images were captured by PhD student Liam Riley working with Professor Smith at the University of Warwick. Prior to microscopy, each sample was treated with a 2% solution of uranyl acetate, which acts as a negative stain. Negative stains work by having negligible interactions with the sample, in this case P(*p*-AT) **S-1**. This leaves the sample surrounded by uranium atoms, which possess a significantly higher electron density to the foldamer.[82,83] When an electron beam is directed at the surface of the sample, the electron beam passes through the foldamer due to its low electron density but not the uranium background. This produces a contrast of a dark background against a light sample, thus enabling the determination of structure and morphology.[82,83] Following staining, the samples could be analysed using transmission electron microscopy (TEM). Initially the 100% DMF solution displayed no visible structures (**Figure 5.18**). This was consistent with our theory of this condition comprised of foldamers with no secondary or tertiary structure; short, ungrouped polymers would not produce enough contrast between the sample and the background.

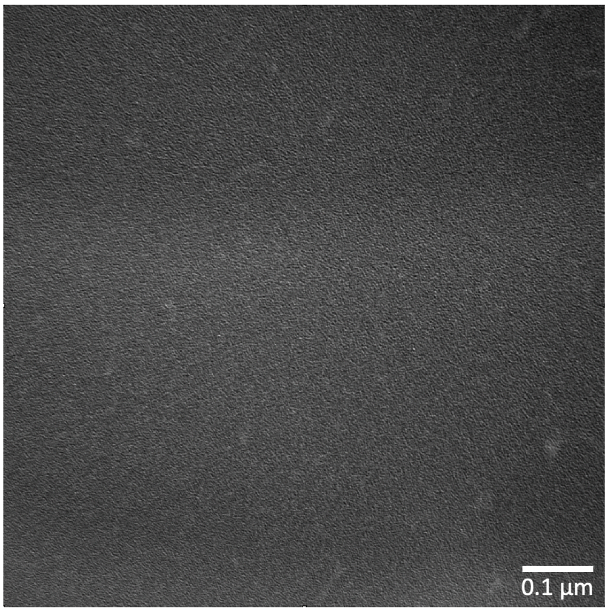


***Figure 5.18**: TEM image captured of P(p-AT) **S-1** in a solvent system composed of 100% DMF. Image captured by Liam Riley from the University of Warwick.*

At 15% water we saw a negative LD signal, indicating that the molecules must be structured enough to align, yet have their **μ** perpendicular to the long axis of the molecules. Upon examination of the TEM (**Figure 5.19**) it is clear to see clusters of small structures (~0.05 µm), further supporting the theory of loose coil formation, which is indicated in all three of the spectroscopic studies. In this small bundle we were able to determine that the **μ** of the monomer is tilted to be less than 54.7° from the long axis of this structure.

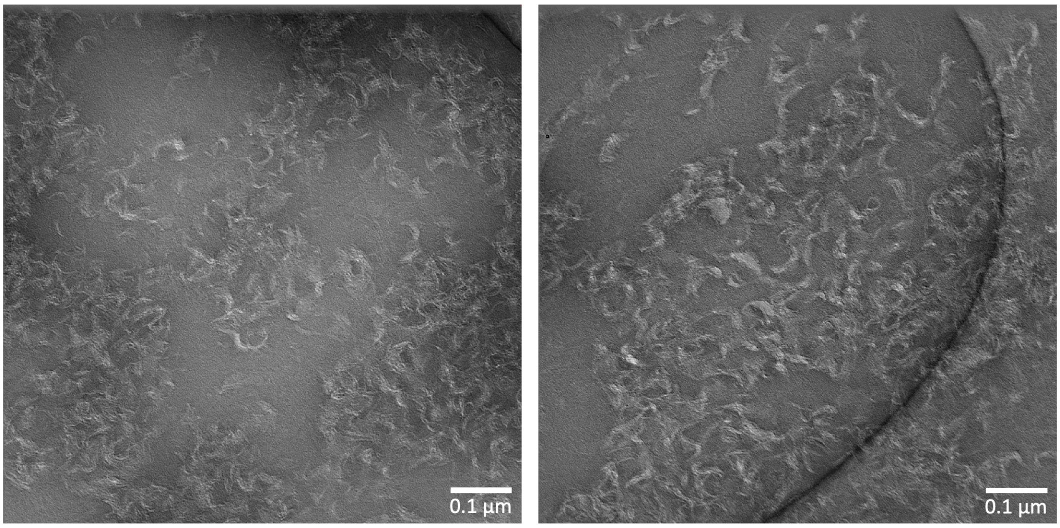


Figure 5.19: *Two TEM images captured of P(p-AT)* ***S-1*** *in a solvent system composed of 85% DMF and 15% water. The images are the same sample taken in two different areas. Image captured by Liam Riley from the University of Warwick.*

The next solvent regime, 45% water was selected as this solvent composition generated the largest LD signal (**Figure 5.14**).

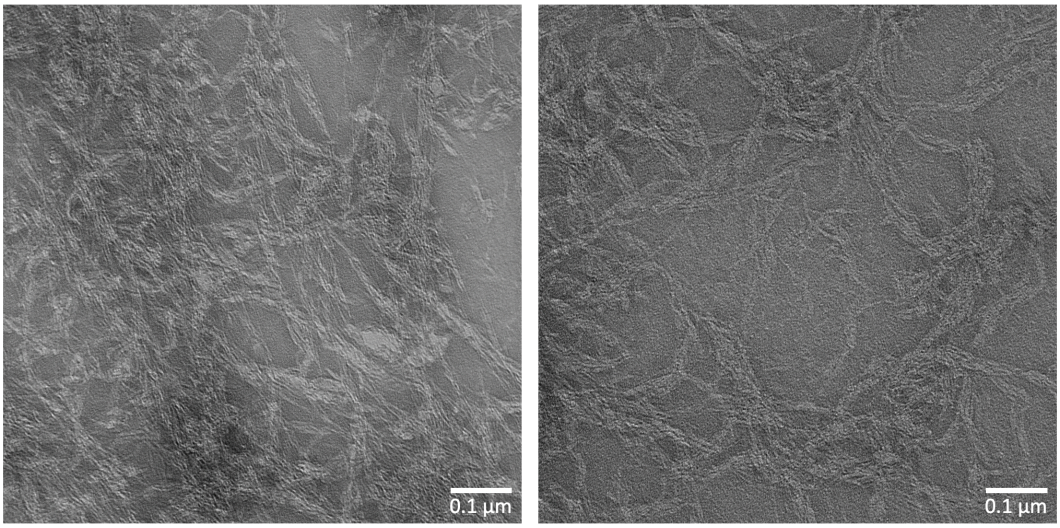


***Figure 5.20**: Two TEM images captured of P(p-AT) **S-1** in a solvent system composed of 55% DMF and 45% water. The images are the same sample taken in two different areas. Image captured by Liam Riley from the University of Warwick.*

This was the condition in which we first questioned the nature of the original assignment - a positive LD signal indicated that the **µ** of the backbone aligned parallel to the alignment axis, which conflicted with the original assumption. Upon analysis of the TEM (**Figure 5.20**), we were able to see long fibrous structures that appeared to be groupings of multiple polymeric chains into single, larger structures. This fitted to our suggestion that the positive LD signal was a result, not of tight coils, but of the formation of a superhelix, in which multiple P(*p*-AT) **S-1** species wrap around one another, rather like homologues to natural molecules such as double stranded DNA or collagen (**Figure 5.21**).[84,85] Upon closer examination of the TEM image, it is clear that each bundle is comprised of individual strands aligned in the same direction as the long axis of the overall bundle; as the individual foldamers twist to form the superhelices, there is a net change in orientation of the aromatic chromophores. As a result, they become more parallel than perpendicular to the flow direction in the LD experiment. This would result in the **µ** of the backbone lining up parallel to the long axis of the superhelix, thereby producing a positive LD signal.

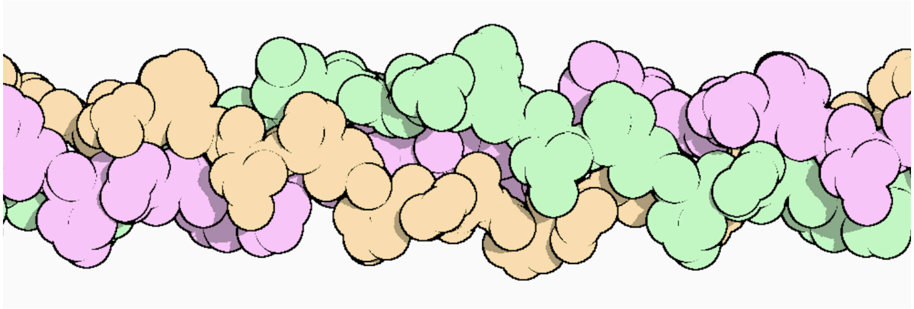


***Figure 5.21**: Triple helix structure of collagen. Image adapted from Goodsell.[84]*

Finally, P(*p*-AT) **S-1** in 90% water was analysed in order to ascertain the potential reason for the decrease in LD signal as the water concentration increases above 45%. The decrease in LD signal implies that the **μ** has shifted away from the long axis of the molecule and approached the so-called 'magic angle' at 54.7°. Examination of the TEM (**Figure 5.22**) quickly revealed the reason behind this - as the water content increases, the majority of the extended structures collapse into spherical aggregates, although some of the superhelices do remain. Using this data, we have determined that the decrease in LD signal is a result of the elimination of a large percentage of the foldamers from solution into the spherical aggregates. As the spherical aggregates do not contain a long axis, they have a low aspect ratio, meaning they tumble when in solution and consequently produce no overall LD signal.

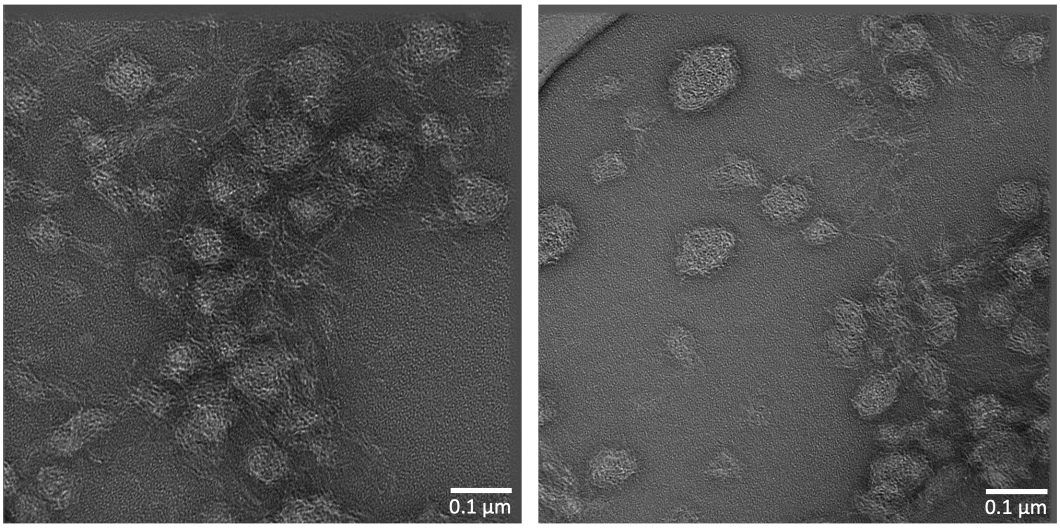


Figure 5.22: *Two TEM images captured of P(p-AT)* ***S-1*** *in a solvent system composed of 10% DMF and 90% water. The images are the same sample taken in two different areas. Image captured by Liam Riley from the University of Warwick.*

5.5 Conclusions

In this study we have investigated the LD activity of self-assembled foldamers, their stability in flow, the orientation of their **μ**, and overall structure. The results of our CD, LD and TEM studies allow us to postulate a revised structure of the P(*p*-AT) **S-1** foldamer assembly compared to that previously put forward in the literature, as follows.[81] Under the conditions of 0 – 9% water content, the foldamer is unfolded and unstructured, consistent with the original findings. As the water increase to 10 – 20% water the structure is a loosely wound helix. The negative LD signal indicates that the monomer **μ** are perpendicular to the orientation axis, while the TEM images provide evidence that supramolecular structures have formed. Meanwhile the positive monosignate CD signal does not give rise to an excitonic CD signal (limited stacking of π systems), but is chiral.[67] As the water increases above >20%, the foldamers adopt a tighter helical confirmation and organise into superhelices, giving rise to a bisignate Cotton effect in CD evident of excitonic coupling; however, at the same time, the TEM images show long structures which are responsible for the large positive LD signal, caused by a change in orientation of the **μ** to be more parallel with the alignment axis. Upon further increase in water content to >55%, there is a decrease in magnitude of LD signal, which the TEM shows are a result of the foldamers having adopted a spherical aggregate formation, which are possibly due to limited water solubility of the foldamers at higher water content. This work shows that the P(*p*-AT) foldamers display a linear structure, implying that they could act as a suitable alternative to M13 as a scaffold.

5.6 Future Work

While we have been able to determine the ability of P(*p*-AT) to generate an LD signal, following which we have been able to investigate the supramolecular assembly of the foldamer, there are a series of studies we would still like to conduct. This will include focusing on the functionalisation of the ethylene glycol side chain to possess a group that is capable of conjugation, potentially a maleimide, which selectively binds to thiols that are easily incorporated onto DNA strands. Furthermore, we would need to conduct stability studies in order to determine whether the foldamer degrades at higher temperatures in the presence of biological materials and with or without covalently conjugated DNA. Following DNA conjugation, the work will focus on the development of an adapted LD-based sandwich assay analogous to the work in Chapter 3, for example with the incorporation of RTF-EXPAR to produce a specific absorbance-based assay for the rapid detection of RNA from viral and other pathogens.

5.7 References

1. Pacheco-Gómez, R. *et al.* Detection of Pathogenic Bacteria Using a Homogeneous Immunoassay Based on Shear Alignment of Virus Particles and Linear Dichroism. *Analytical Chemistry* **84**, 91–97 (2012).
2. Carr-Smith, J. *et al.* Polymerase Chain Reaction on a Viral Nanoparticle. *ACS Synthetic Biology* **4**, 1316–1325 (2015).
3. Tridgett, M. *et al.* Linear dichroism of visible-region chromophores using M13 bacteriophage as an alignment scaffold. *RSC Advances* **8**, 29535–29543 (2018).
4. Little, H. A. The Development of Novel Diagnostic Sensors Based on Linear Dichroism Spectroscopy. (University of Birmingham, 2016).
5. Smeal, S. W., Schmitt, M. A., Pereira, R. R., Prasad, A. & Fisk, J. D. Simulation of the M13 life cycle II: Investigation of the control mechanisms of M13 infection and establishment of the carrier state. *Virology* **500**, 275–284 (2017).
6. Mandelkern, L. Macromolecules of Biological Importance. in *An Introduction to Macromolecules* (ed. Mandelkern, L.) 100–152 (Springer New York, 1983). doi:10.1007/978-1-4612-5494-2_6.
7. Gellman, S. H. Foldamers: A Manifesto. *Acc. Chem. Res.* **31**, 173–180 (1998).
8. Hill, D. J., Mio, M. J., Prince, R. B., Hughes, T. S. & Moore, J. S. A Field Guide to Foldamers. *Chem. Rev.* **101**, 3893–4012 (2001).
9. Jeffrey, G. A. & Saenger, W. *Hydrogen Bonding in Biological Structures*. (Springer Berlin Heidelberg, 1991).
10. Le Bailly, B. A. F. & Clayden, J. Dynamic foldamer chemistry. *Chem. Commun.* **52**, 4852–4863 (2016).
11. Hecht, S. & Huc, I. *Foldamers: Structure, Properties and Applications*. (John Wiley & Sons, 2007).
12. Ikkanda, B. A. & Iverson, B. L. Exploiting the interactions of aromatic units for folding and assembly in aqueous environments. *Chem. Commun.* **52**, 7752–7759 (2016).

13. Huc, I. Aromatic Oligoamide Foldamers. *Eur. J. Org. Chem.* **2004**, 17–29 (2004).
14. Xin, P., Zhu, P., Su, P., Hou, J.-L. & Li, Z.-T. Hydrogen-Bonded Helical Hydrazide Oligomers and Polymer That Mimic the Ion Transport of Gramicidin A. *J. Am. Chem. Soc.* **136**, 13078–13081 (2014).
15. Ferrand, Y. *et al.* Diastereoselective Encapsulation of Tartaric Acid by a Helical Aromatic Oligoamide. *J. Am. Chem. Soc.* **132**, 7858–7859 (2010).
16. Juwarker, H., Suk, J. & Jeong, K.-S. Foldamers with helical cavities for binding complementary guests. *Chem. Soc. Rev.* **38**, 3316 (2009).
17. Lang, C. *et al.* Biomimetic Transmembrane Channels with High Stability and Transporting Efficiency from Helically Folded Macromolecules. *Angew. Chem. Int. Ed.* **55**, 9723–9727 (2016).
18. Pfukwa, R., Kouwer, P. H. J., Rowan, A. E. & Klumperman, B. Templated Hierarchical Self-Assembly of Poly(p-aryltriazole) Foldamers. *Angew. Chem. Int. Ed.* **52**, 11040–11044 (2013).
19. Ousaka, N., Yamaguchi, T. & Yashima, E. Remarkable Enhancement of Stability and Helix-sense Excess of Oligo(phenylene ethynylene) Foldamers Assisted by Linking with Achiral (Metallo)salen Tethers and Their Application to Asymmetric Catalysis. *Chem. Lett.* **43**, 512–514 (2014).
20. Zheng, L. *et al.* Chiral bisphosphine ligands based on quinoline oligoamide foldamers: application in asymmetric hydrogenation. *Org. Biomol. Chem.* **17**, 9573–9577 (2019).
21. Zega, A. Azapeptides as Pharmacological Agents. *CMC* **12**, 589–597 (2005).
22. Jiménez Blanco, J. L., Bootello, P., Benito, J. M., Ortiz Mellet, C. & García Fernández, J. M. Urea-, Thiourea-, and Guanidine-Linked Glycooligomers as Phosphate Binders in Water. *J. Org. Chem.* **71**, 5136–5143 (2006).
23. Kawauchi, T., Kitaura, A., Kumaki, J., Kusanagi, H. & Yashima, E. Helix-Sense-Controlled Synthesis of Optically Active Poly(methyl methacrylate) Stereocomplexes. *J. Am. Chem. Soc.* **130**, 11889–11891 (2008).

24. Hamuro, Y., Geib, S. J. & Hamilton, A. D. Novel Molecular Scaffolds: Formation of Helical Secondary Structure in a Family of Oligoanthranilamides. *Angew. Chem. Int. Ed. Engl.* **33**, 446–448 (1994).
25. Hamuro, Y., Geib, S. J. & Hamilton, A. D. Oligoanthranilamides. Non-Peptide Subunits That Show Formation of Specific Secondary Structure. *J. Am. Chem. Soc.* **118**, 7529–7541 (1996).
26. Gothard, C. M., Rao, N. A. & Nowick, J. S. Nanometer-Sized Amino Acids for the Synthesis of Nanometer-Scale Water-Soluble Molecular Rods of Precise Length. *J. Am. Chem. Soc.* **129**, 7272–7273 (2007).
27. Rodriguez, J. M. & Hamilton, A. D. Benzoylurea Oligomers: Synthetic Foldamers That Mimic Extended α Helices. *Angew. Chem. Int. Ed.* **46**, 8614–8617 (2007).
28. Zhang, A., Han, Y., Yamato, K., Zeng, X. C. & Gong, B. Aromatic Oligoureas: Enforced Folding and Assisted Cyclization. *Org. Lett.* **8**, 803–806 (2006).
29. Clayden, J., Lemiègre, L. & Helliwell, M. Synthesis and Stacked Conformations of Symmetrical and Unsymmetrical Oligo-ureas of Metaphenylenediamine. *J. Org. Chem.* **72**, 2302–2308 (2007).
30. Clayden, J., Lemiègre, L., Pickworth, M. & Jones, L. Conformation and stereodynamics of 2,2′-disubstituted N,N′-diaryl ureas. *Org. Biomol. Chem.* **6**, 2908 (2008).
31. Liu, R., Connor, A. L., Al-mkhaizim, F. Y. & Gong, B. Aromatic oligoamides with increased backbone flexibility: improved synthetic efficiencies, solvent-dependent folding and cooperative conformational transitions. *New J. Chem.* **39**, 3217–3220 (2015).
32. Kudo, M. & Tanatani, A. Conformational properties of aromatic multi-layered and helical oligoureas and oligoguanidines. *New J. Chem.* **39**, 3190–3196 (2015).
33. Ben, T., Furusho, Y., Goto, H., Miwa, K. & Yashima, E. Double helix formation of poly(m-phenylene)s bearing achiral oligo(ethylene oxide) pendants and transformation into an excess of one-handed single helix through cholate binding in water. *Org. Biomol. Chem.* **7**, 2509 (2009).

34. Meudtner, R. M. & Hecht, S. Helicity Inversion in Responsive Foldamers Induced by Achiral Halide ion Guests. *Angew. Chem. Int. Ed.* **47**, 4926–4930 (2008).
35. Peters, A. D. *et al.* Switchable foldamer ion channels with antibacterial activity. *Chem. Sci.* **11**, 7023–7030 (2020).
36. Lister, F. G. A. *et al.* Bis-pyrene probes of foldamer conformation in solution and in phospholipid bilayers. *Chem. Sci.* **9**, 6860–6870 (2018).
37. Jeon, H.-G., Jung, J. Y., Kang, P., Choi, M.-G. & Jeong, K.-S. Folding-Generated Molecular Tubes Containing One-Dimensional Water Chains. *J. Am. Chem. Soc.* **138**, 92–95 (2016).
38. Meunier, A. *et al.* Aromatic foldamers as scaffolds for metal second coordination sphere design. *Chem. Sci.* **11**, 12178–12186 (2020).
39. Miyoshi, T. Characterization of polymers by NMR. in *Molecular Characterization of Polymers* 409–440 (Elsevier, 2021). doi:10.1016/B978-0-12-819768-4.00013-0.
40. Kan, R. O. A Correlation of Chemical Shifts with Inductive Effect Parameters. *J. Am. Chem. Soc.* **86**, 5180–5183 (1964).
41. Abragam, A. *The principles of nuclear magnetism*. (Oxford Univ. Pr, 2011).
42. Nelson, J. C. Solvophobically Driven Folding of Nonbiological Oligomers. *Science* **277**, 1793–1796 (1997).
43. Steinwand, S., Yu, Z., Hecht, S. & Wachtveitl, J. Ultrafast Dynamics of Photoisomerization and Subsequent Unfolding of an Oligoazobenzene Foldamer. *J. Am. Chem. Soc.* **138**, 12997–13005 (2016).
44. Lu, Z., Zhu, Y., Lin, J., Jiang, X. & Li, Z. Hydrogen bonded foldamer-bridged biscoumarins: A UV-Vis absorption and fluorescent study of the solvent effect. *Chin. Sci. Bull.* **55**, 2870–2878 (2010).
45. Rhodes, W. Hypochromism and Other Spectral Properties of Helical Polynucleotides. *J. Am. Chem. Soc.* **83**, 3609–3617 (1961).
46. Tinoco, I. Hypochromism in Polynucleotides 1. *J. Am. Chem. Soc.* **82**, 4785–4790 (1960).

47. Kudo, M., Maurizot, V., Masu, H., Tanatani, A. & Huc, I. Structural elucidation of foldamers with no long range conformational order. *Chem. Commun.* **50**, 10090–10093 (2014).
48. Carini, M. *et al.* High conductance values in π-folded molecular junctions. *Nat Commun* **8**, 15195 (2017).
49. Gin, M. S., Yokozawa, T., Prince, R. B. & Moore, J. S. Helical Bias in Solvophobically Folded Oligo(Phenylene Ethynylene)s. *J. Am. Chem. Soc.* **121**, 2643–2644 (1999).
50. Brunsveld, L., Zhang, H., Glasbeek, M., Vekemans, J. A. J. M. & Meijer, E. W. Hierarchical Growth of Chiral Self-Assembled Structures in Protic Media †. *J. Am. Chem. Soc.* **122**, 6175–6182 (2000).
51. Cai, W. *et al.* Foldamer Organogels: A Circular Dichroism Study of Glucose-Mediated Dynamic Helicity Induction and Amplification. *J. Am. Chem. Soc.* **130**, 13450–13459 (2008).
52. Prince, R. B., Brunsveld, L., Meijer, E. W. & Moore, J. S. Twist Sense Bias Induced by Chiral Side Chains in Helically Folded Oligomers. *Angewandte Chemie International Edition* **39**, 228–230 (2000).
53. Prince, R. B., Moore, J. S., Brunsveld, L. & Meijer, E. W. Cooperativity in the Folding of Helical m-Phenylene Ethynylene Oligomers Based upon the `Sergeants-and-Soldiers’ Principle. *Chemistry – A European Journal* **7**, 4150–4154 (2001).
54. Grime, R. L. *et al.* Differences in SMA-like polymer architecture dictate the conformational changes exhibited by the membrane protein rhodopsin encapsulated in lipid nano-particles. *Nanoscale* **13**, 13519–13528 (2021).
55. Ball, L. E. *et al.* Influence of DIBMA Polymer Length on Lipid Nanodisc Formation and Membrane Protein Extraction. *Biomacromolecules* **22**, 763–772 (2021).
56. Stroud, Z., Hall, S. C. L. & Dafforn, T. R. Purification of membrane proteins free from conventional detergents: SMA, new polymers, new opportunities and new insights. *Methods* **147**, 106–117 (2018).

57. Grime, R. L. *et al.* Single molecule binding of a ligand to a G-protein-coupled receptor in real time using fluorescence correlation spectroscopy, rendered possible by nano-encapsulation in styrene maleic acid lipid particles. *Nanoscale* **12**, 11518–11525 (2020).
58. Creager, A. N. H., Scholthof, K.-B. G., Citovsky, V. & Scholthof, H. B. Tobacco Mosaic Virus: Pioneering Research for a Century. *Plant Cell* **11**, 301–308 (1999).
59. Clack, B. A. & Gray, D. M. Flow linear dichroism spectra of four filamentous bacteriophages: DNA and coat protein contributions. *Biopolymers* **32**, 795–810 (1992).
60. Rodger, A. How to Study DNA and Proteins by Linear Dichroism Spectroscopy. *Science Progress* **91**, 377–396 (2008).
61. Niu, Z., Bruckman, M. A., Harp, B., Mello, C. M. & Wang, Q. Bacteriophage M13 as a scaffold for preparing conductive polymeric composite fibers. *Nano Res.* **1**, 235–241 (2008).
62. Liu, X. *et al.* Size Dependent Cellular Uptake of Rod-like Bionanoparticles with Different Aspect Ratios. *Sci Rep* **6**, 24567 (2016).
63. Teo, A. C. K. *et al.* Analysis of SMALP co-extracted phospholipids shows distinct membrane environments for three classes of bacterial membrane protein. *Sci Rep* **9**, 1813 (2019).
64. The C=O Bond, Part III: Carboxylic Acids. *Spectroscopy Online* https://www.spectroscopyonline.com/view/co-bond-part-iii-carboxylic-acids.
65. The C=O Bond, Part VI: Esters and the Rule of Three. *Spectroscopy Online* https://www.spectroscopyonline.com/view/co-bond-part-vi-esters-and-rule-three.
66. Razmkhah, K., Chmel, N. P., Gibson, M. I. & Rodger, A. Oxidized polyethylene films for orienting polar molecules for linear dichroism spectroscopy. *Analyst* **139**, 1372–1382 (2014).
67. Nordén, B., Rodger, A. & Dafforn, T. *Linear dichroism and circular dichroism: a textbook on polarized-light spectroscopy*. (Royal Society of Chemistry, 2010).

68. Aflak, N. *et al.* Facile immobilization of copper(I) acetate on silica: A recyclable and reusable heterogeneous catalyst for azide–alkyne clickable cycloaddition reactions. *Polyhedron* **170**, 630–638 (2019).
69. Joshi, S. M. *et al.* Synthesis of radiolabelled aryl azides from diazonium salts: experimental and computational results permit the identification of the preferred mechanism. *Chem. Commun.* **51**, 8954–8957 (2015).
70. Breugst, M. & Reissig, H. The Huisgen Reaction: Milestones of the 1,3-Dipolar Cycloaddition. *Angew. Chem. Int. Ed.* **59**, 12293–12307 (2020).
71. Elliott, P. I. P. Chapter 1. Organometallic complexes with 1,2,3-triazole-derived ligands. in *Organometallic Chemistry* (eds. Fairlamb, I. J. S. & Lynam, J. M.) vol. 39 1–25 (Royal Society of Chemistry, 2014).
72. Jiang, Y., Kuang, C. & Yang, Q. The Use of Calcium Carbide in the Synthesis of 1-Monosubstituted Aryl 1,2,3-Triazole via Click Chemistry. *Synlett* **2009**, 3163–3166 (2009).
73. Sonogashira, K. Development of Pd–Cu catalyzed cross-coupling of terminal acetylenes with sp2-carbon halides. *Journal of Organometallic Chemistry* **653**, 46–49 (2002).
74. Yin & Liebscher, J. Carbon−Carbon Coupling Reactions Catalyzed by Heterogeneous Palladium Catalysts. *Chem. Rev.* **107**, 133–173 (2007).
75. Sonogashira Coupling. *Chemistry LibreTexts* https://chem.libretexts.org/Bookshelves/Inorganic_Chemistry/Modules_and_Websites_(Inorganic_Chemistry)/Catalysis/Catalyst_Examples/Sonogashira_Coupling (2016).
76. Henkes, L. M., Haus, P., Jäger, F., Ludwig, J. & Meyer-Almes, F.-J. Synthesis and biochemical analysis of 2,2,3,3,4,4,5,5,6,6,7,7-dodecafluoro-N-hydroxy-octanediamides as inhibitors of human histone deacetylases. *Bioorganic & Medicinal Chemistry* **20**, 985–995 (2012).
77. Marrington, R. *et al.* Validation of new microvolume Couette flow linear dichroism cells. *The Analyst* **130**, 1608 (2005).

78. Matsuoka, Y. & Norden, B. Linear dichroism studies of nucleic acid bases in stretched poly(vinyl alcohol) film. Molecular orientation and electronic transition moment directions. *J. Phys. Chem.* **86**, 1378–1386 (1982).

79. McLachlan, J. R. A., Smith, D. J., Chmel, N. P. & Rodger, A. Calculations of flow-induced orientation distributions for analysis of linear dichroism spectroscopy. *Soft Matter* **9**, 4977 (2013).

80. Bulheller, B. M., Rodger, A. & Hirst, J. D. Circular and linear dichroism of proteins. *Physical Chemistry Chemical Physics* **9**, 2020 (2007).

81. Carter, J. G. *et al.* Linear Dichroism Activity of Chiral Poly(p-Aryltriazole) Foldamers. *ACS Omega* acsomega.1c06139 (2021) doi:10.1021/acsomega.1c06139.

82. De Carlo, S. & Harris, J. R. Negative staining and cryo-negative staining of macromolecules and viruses for TEM. *Micron* **42**, 117–131 (2011).

83. Hou, J. Electron microscopy for tight junction. in *A Laboratory Guide to the Tight Junction* 213–261 (Elsevier, 2020). doi:10.1016/B978-0-12-818647-3.00006-4.

84. PDB101: Molecule of the Month: Collagen. *RCSB: PDB-101* http://pdb101.rcsb.org/motm/4.

85. Dafforn, T. R., Della, M. & Miller, A. D. The Molecular Interactions of Heat Shock Protein 47 (Hsp47) and Their Implications for Collagen Biosynthesis. *Journal of Biological Chemistry* **276**, 49310–49319 (2001).

Chapter 6: Experimental

6.1 Materials and methods

All reagents and solvents were purchased from Sigma Aldrich (Gillingham, UK), New England Biolabs (Hitchin, UK) or Thermo Fisher Scientific (Loughborough, UK) and used without further purification unless otherwise stated. Commercial LAMP and PCR kits used in Chapter 4 were purchased from New England Biolabs (Hitchin, UK) and Pro-lab Diagnostics (Birkenhead, UK) respectively. Foldamers analysed in Chapter 5 were synthesised and supplied by Dr Rueben Pfukwa from Stellenbosch University (Stellenbosch, ZA). DNA was purchased from Sigma Aldrich (Gillingham, UK) and RNA was supplied by Public Health England (PHE). Milli-Q water purified with a Millipore Elix-Gradient A10 system (resistivity > 18µΩ.cm, TOC ≤ 5ppb, Millipore, France) was used for DNA sample preparation and in DNA and RNA containing experiments.

6.2 Data Analysis

All spectra was examined and plotted using the data analysis and graphing software application GraphPad Prism 9.

6.3 Chapter 3 Experimental

6.3.1 Media and Buffers

6.3.1.1 Poly(ethylene glycol) solution

250 g of Poly(ethylene glycol) (PEG; MW 6000) and 146.1 g sodium chloride (NaCl) were combined in 1 L of distilled water. The resulting solution was sterilised by autoclave at 121 °C for 20 minutes.

6.3.1.2 Potassium Phosphate Buffers

6.3.1.2.1 1 M dipotassium phosphate solution (K_2HPO_4)

17.418 g of dipotassium phosphate was dissolved in 100 mL of distilled water.

6.3.1.2.2 1 M potassium dihydrogen phosphate solution (KH_2PO_4)

13.609 g of potassium dihydrogen phosphate (KH_2PO_4) was dissolved in 100 mL of distilled water.

6.3.1.2.3 50 mM potassium phosphate buffer pH 8.0

47 mL of 1 M K_2HPO_4 (Chapter 6.3.1.2.1) and 3 mL of 1 M KH_2PO_4 (Chapter 6.3.1.2.2) were diluted with 950 mL of distilled water. The pH of the buffer was checked using a pH meter and adjusted to pH 8.0 if needed. The resulting solution was sterilised by autoclave at 121 °C for 20 minutes.

6.3.1.3 100 mM potassium phosphate buffer pH 7.6

86.6 mL of 1 M K_2HPO_4 (Chapter 6.3.1.2.1) and 13.4 mL of 1 M KH_2PO_4 (Chapter 6.3.1.2.2) were diluted with 900 mL of distilled water. The pH of the buffer was checked using a pH meter and adjusted to pH 7.6 if needed. The resulting solution was sterilised by autoclave at 121 °C for 20 minutes.

6.3.1.4 10x Phosphate Buffered Saline (PBS)

17.8 g of disodium phosphate (Na_2HPO_4), 2.4 g of potassium dihydrogen phosphate (KH_2PO_4), 80 g of NaCl and 2 g of potassium chloride (KCl) were combined in 1 L of distilled water. The resulting solution was sterilised by autoclave at 121 °C for 20 minutes.

6.3.1.4.1 1x Phosphate Buffered Saline (PBS)

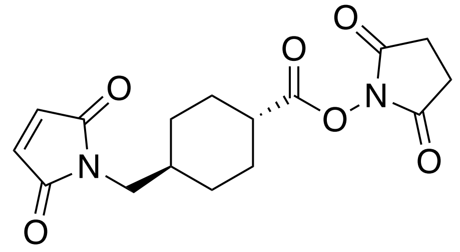
100 mL of 10x PBS solution was combined with 900 mL of distilled water. The pH of the buffer was checked using a pH meter and adjusted to pH 7.4 if required. The resulting solution was sterilised by autoclave at 121 °C for 20 minutes.

6.3.2 M13 Bacteriophage Propagation

Six 2 L conical flasks containing 400 mL of 25 g/L Nutrient Broth 2 (NB2) (Oxoid) were sterilised via autoclave at 121 °C for 20 minutes. Under sterile conditions the media solutions were inoculated with 400 µL of 5 mg/mL of tetracycline dissolved in ethanol and 500 µL of One Shot Top10F´ *E. coli* (Invitrogen, California, USA) doped with 10 µL of 10 mg/mL M13 bacteriophage (New England Biolabs, UK). The culture was incubated at 37 °C, at 200 rpm for 16 hours in an orbital shaker. Following incubation, the culture was centrifuged (Avanti J-24, Beckman Coulter USA centrifuge with a JLA 10.5 rotor) at 8500 rpm for 30 minutes at 4 °C to pellet the bacterial cells. To ensure all the *E. coli* had been removed the supernatant was taken and centrifuged under the same conditions. Following the second centrifugation the top 80% of the supernatant was collected and the M13 precipitated using PEG solution (Chapter 6.3.1.1) in 5:1 ratio (Supernatant/PEG). The solution was allowed to stir at 4 °C for 90 minutes before being centrifuged (Avanti J-24, Beckman Coulter, USA centrifuge with a JLA 10.5 rotor) at 10000 rpm for 25 minutes at 4 °C. The supernatant was discarded, and each pellet resuspended in 1 mL of 50 mM potassium phosphate buffer pH 8.0 (Chapter 6.3.1.2.3) in a microcentrifuge tube. To remove any residual *E. coli* the tubes were centrifuged (FX241.5p rotor, Microfuge® 16, Beckman Coulter, USA) at 14000 rpm for 5 minutes. The pellet was discarded and the M13 was precipitated from the supernatant with PEG solution (a ratio of 5:1 supernatant/PEG). The solution was placed in ice for 1 hour. The solution was centrifuged (FX241.5p rotor, Microfuge® 16, Beckman Coulter, USA) at 14000 rpm for 15 minutes. The resulting supernatant was discarded, and each pellet was resuspended in 200 µL of 100 mM potassium phosphate buffer pH 8.0. The solutions were pooled and stored at 4 °C.[1]

6.3.3 Synthesis

6.3.3.1 Succinimidyl 4-(N-maleimidomethyl) cyclohexane–1–carboxylate, SMCC



A known compound synthesised according to a literature procedure.[2]

(Flask A) Trans-4-(aminomethyl)cyclohexane carboxylic acid (0.79 g, 5 mmol) and maleic anhydride (0.49 g, 5 mmol) were suspended in DMF (25 mL) and stirred at room temperature for 6 hours under argon. The solution was cooled to 0 °C and 2,4,6-trimethylpyridine (1.39 mL, 10.5 mmol) was added dropwise. In a separate flask (Flask B), a solution of *N*-hydroxysuccinimide (2.30 g, 20 mmol) in DMF (25 mL) was stirred at 0 °C under argon as trifluoroacetic anhydride (2.78 mL, 20 mmol) was added dropwise. The reaction mixture was stirred for 10 min, and 2,4,6-trimethylpyridine (2.64 mL, 20 mmol) was added dropwise and stirred for a further 10 minutes. The solution in Flask B was added by syringe to Flask A dropwise. After addition was complete, the reaction mixture warmed to room temperature overnight. The reaction mixture was diluted with CH_2Cl_2 (30 mL) and 1 M HCl (25 mL). The organic layer was washed with 1 M HCl (2 x 25 mL). The combined organic layers were dried with $MgSO_4$, filtered, and concentrated *in vacuo* to afford an orange solid. The solid was triturated with Et_2O (3 x 20 mL) to afford **SMCC** as a white solid (1.64 g, 98%). **m.p.** = 173- 176 °C; **^{1}H NMR** 6.71 (s, 2H), 3.39 (d, *J* = 7.0 Hz, 2H), 2.82 (s, 4H), 2.58 (tt, *J* = 12.2, 3.6 Hz, 1H), 2.16 (d, *J* = 10.5 Hz, 1H), 1.54 (dt, *J* = 13.3, 9.9 Hz, 3H), 1.06 (qd, *J* = 13.3, 3.5 Hz, 2H);**^{13}C NMR** (101 MHz, DMSO-*d*$_6$): δ (ppm) = 170.99, 170.60, 169.14, 134.04, 43.41, 40.42, 36.09, 29.32, 28.02, 25.60; **HRMS** (ESI): *m/z* calcd for $C_{16}H_{18}N_2O_6$: 334.33 [M]; found: 357.11 [M + Na].

Analytical data in agreement with literature values.[2]

6.3.4 M13 Bacteriophage Bioconjugation

Oligonucleotides conjugated to M13 bacteriophage were synthesised with a 5´ or 3´ thiol linked group as a disulfide. The disulfide was reduced prior to use conjugation through the addition of 5 mM TCEP·HCl (5:1 molar ration TCEP/DNA) and by allowing to mix for 1 hour at RT. To a 1.5 mL microcentrifuge tube was added, 1 mL of M13 Bacteriophage (2 mg/mL in 50 mM potassium phosphate buffer, pH 8.0) and 88 µL of SMCC solution (24 mg/mL in DMSO) and the solution mixed for 1 hour at 25 °C. Unreacted SMCC was removed from the mixture using a PD-10 G-25 column (GE Healthcare), the conjugated M13 was eluted into 3.5 mL of 100 mM potassium phosphate buffer, pH 7.6. The M13 was combined with the reduced oligonucleotides and mixed at RT for 16 hours. 110 µL of L-cysteine (10 mg/mL in water) was added to the M13-DNA solution, to quench any unreacted SMCC, and allowed to mix at RT for 1 hour. The M13-DNA was precipitated from solution through the addition of PEG (3:1 ratio M13-DNA solution/PEG) and incubation at 4 °C in the dark for 1 hour. The solution was centrifuged (FX241.5p rotor, Microfuge® 16, Beckman Coulter, USA) at 14000 rpm for 10 minutes. The supernatant was discarded, and the pellet resuspended in 1 mL of 1x PBS (Chapter 6.3.1.4.1), before being stored at 4 °C.

6.3.5 UV-Vis Spectroscopy

All measurements were performed on a Jasco V-750 UV-Visible spectrophotometer (Jasco, Japan) in a 1 mL quartz cuvette. Using the following parameters:

Table 6.1: *Table of the parameters used during UV-vis measurements.*

Wavelength Range	200 – 400 nm
Scan Speed	50 nm/min
Data Pitch	0.1 nm
Bandwidth	1 nm

6.3.5.1 M13 Bacteriophage Concentration Determination

A blank was recorded of 1000 µL of 50 mM potassium phosphate buffer pH 8.0. Samples were recorded of solutions containing 10 µL of M13 Bacteriophage mixed with 990 µL of 50 mM potassium phosphate buffer pH 8.0. The blank was subtracted from the sample and zeroed at 400 nm. The absorbance at wavelength 269 nm was recorded and inserted into **Equation 6.1.** Where A_{269} is the absorbance at 269 nm, Df is the dilution factor of 100, ε is the extinction coefficient of M13 bacteriophage of 3.84 cm^2/mg, and l is the path length of 1 cm. The calculated concentration is in mg/mL.

$$C = \frac{A_{269} \times Df}{\varepsilon l}$$

Equation 6.1*: The Beer-Lambert Law rearranged in order to calculate the concentration of M13 bacteriophage.*

6.3.5.2 Oligonucleotide Concentration Determination

A blank was recorded of 1000 µL of water. Samples recorded from solutions composed of 20 µL of DNA stock in 980 µL of water. The baseline was subtracted from the sample and zeroed at 400 nm. The absorbance at wavelength 260 nm was recorded and inserted into **Equation 6.2,** where A_{260} is the absorbance at 260 nm, *Df* is the dilution factor of 50, ε is the extinction coefficient of the oligonucleotide being analysed (**Table 3.3**), and *l* is the path length of 1 cm. The calculated concentration is in µM.

$$C = \frac{A_{260} \times Df}{\varepsilon l}$$

Equation 6.2*: The Beer-Lambert Law rearranged in order to calculate the concentration of oligonucleotides.*

6.3.6 Exponential Amplification Reaction

The protocol first involves the preparation of three solutions, Part A, Part B and Part C, followed by an addition step and then finally an amplification step.

6.3.6.1 Part A

0.90 µL of water, 2.50 µL of 10x Isothermal amplification buffer, 3.75 µL of BSA solution, 1.50 µL of *Bst* 2.0 DNA polymerase (1.6 U/µL), 1.35 µL of Nt.*Bst*NBI (10 U/µL).

6.3.6.2 Part B

2.85 µL of water, 2.5 µL of 10x Isothermal amplification buffer, 2.40 µL of $MgSO_4$ (100 mM), 0.75 µL dNTP (10 nM), 1.50 µL of dsGreen (1:5 dilution in DMSO from 100x to 20x) (Lumiprobe, De).

6.3.6.3 Part C

4.50 µL of water, 2.5 µL of 10x Isothermal amplification buffer, 1.5 µL of **Template *Chlamydia*´-*Chlamydia*´** (1 µM), 1.50 µL of SSB solution, 3 µL of ***Chlamydia* Trigger** (100 nM, 10 nM, 1 nM, 100 pM, 10 pM, 1 pM or a blank).

OR

1.50 µL of water, 2.5 µL of 10x Isothermal amplification buffer, 1.5 µL of **Template *Chlamydia*´-*Chlamydia*´** (1 µM), 1.50 µL of SSB solution, 3 µL of **Template *Chlamydia*´-Reporter´** (4 µM), 3 µL of ***Chlamydia* Trigger** (100 nM, 10 nM, 1 nM, 100 pM, 10 pM, 1 pM or a blank).

6.3.6.4 Addition Step

Part B (10 µL) is added to a PCR tube, and to this is added Part C (13 µL), followed by Part A (10 µL). The tube is then sealed, with the contents then subjected to amplification.

6.3.6.5 EXPAR Amplification

Isothermal incubation and fluorescence signal measurements were carried out using an Agilent AriaMX Real-Time PCR system (Didcot, UK) set to a constant temperature of 50 °C. The fluorescence is measured every 10 seconds over an incubation time of 15 minutes

6.3.7 Reversed-Phase - High Performance Liquid Chromatography (RP-HPLC)

All analytical RP-HPLC studies were performed using an Agilent 1260 infinity II LC system (Stockport, UK). The column was a Phenomenex (Macclesfield, UK) Clarity Oligo-Reverse Phase, incorporating C18 packing with a 5 μm particle size, 110 Å pore size, and 250 x 4.6 mm column dimensions. Buffer A is 100% degassed acetonitrile and buffer D is 0.1 M triethylammonium acetate (TEAA). The column was heated to 60 °C prior to sample injection. Samples (20 μL) were injected with a run time of 45 minutes for each sample, at a flow rate of 1ml/min. The solvent gradients used can be found in **Table 6.2**. The absorbance of each run was monitored at 260nm.

***Table 6.2**: Solvent gradient used for analytic RP-HPLC.*

Time (Minutes)	Buffer A (%)	Buffer D (%)
0	5	95
30	18	82
40	100	0
40.1	5	95
45	5	95

6.3.8 Mass Spectrometry

All mass spectrometry of DNA and amplification solution were performed using a Waters Xevo G2-XS Time of flight instrument (Elstree, UK) incorporating a Waters ACQUITY UPLC Oligonucleotide column (C18 packing, 1.7 μm particle size, 2.1 mm x 50 mm). Buffer A was composed of 75 mM Triethylammonium acetate (TEAA) in H_2O. Buffer B was composed of 75 mM TEAA in acetonitrile.

6.3.8.1 Analysis of EXPAR Products

10 μL samples of EXPAR solution post incubation (Chapter 6.3.6) were injected into the LC-MS and separated using the conditions stated in **Table 6.2**. MS spectra were obtained using electrospray negative ion mode ranging from 400 to 6000 atomic mass units.

6.3.9 Linear Dichroism Spectroscopy

All measurements were performed on a Jasco J-1500 spectropolarimeter (Jasco, Japan) which has been modified to measure LD.

6.3.9.1 Measurement of LD spectra using Couette Flow Alignment

Samples (100 µL) were placed into a micro-Couette cell.[3] Baselines were recorded from a non-rotating cell; samples were then recorded from a rotating cell (3000 rpm) the baseline was subtracted and zeroed at 400 nm.

***Table 6.3**: Table of the parameters using during Couette flow LD measurements.*

Wavelength Range	200 – 400 nm
Scan Speed	100 nm/min
Data Pitch	0.1 nm
Bandwidth	1.0 nm
Response Time	1 second
Accumulations	3
Scanning mode	Continuous

6.3.9.2 Linear Dichroism Analysis of M13 – DNA Probes

25 µL of M13-DNA was mixed with 75 µL of 1x PBS. The sample was analysed in concordance with Chapter 6.3.9.1. The LD value at 225 nm was taken and inserted into **Equation 6.3** to calculate the volume of M13-DNA probe required to obtain an LD value of 0.1 at 225 nm in 100 µL.

$$Volume\ of\ Probe = \frac{100}{\left(\frac{LD_{225} \times 4}{0.01}\right)}$$

Equation 6.3: The equation used in order to determine the quantity of M13-DNA probe required in order to achieve and LD value of 0.01 at 225 nm.

6.3.9.3 Linear Dichroism Analysis of DNA Targets

To detect target DNA using M13 probes, the volume of probe calculated in Equation 6.3 was combined with 10 µL of analyte (**Table 6.4**) and made up to 100 µL using 1x PBS. The resulting solution was placed into a micro-Couette cell and analysed in concordance with the preceding method (Chapter 6.3.9.1).

***Table 6.4**: Table of target analytes detected by M13 probes.*

Analyte	Use	Concentration Range
Water	Blank for M13 Quantification	N/A
Reporter in water	M13 Quantification	0 – 100 nM (10 nM intervals)
***Chlamydia* Trigger** in water	M13 Quantification	100 nM
EXPAR solution *sans* enzymes	EXPAR Blank	N/A
EXPAR Solution post incubation	EXPAR Analysis	N/A

6.4 Chapter 4 Experimental

6.4.1 UV – Vis Spectroscopy

All measurements were performed on a Jasco V-750 UV-Visible spectrophotometer (Jasco, Japan) in a 1 mL quartz cuvette. Using the following parameters:

Table 6.5: *Table of the parameters used during UV-vis measurements.*

Wavelength Range	200 – 400 nm
Scan Speed	50 nm/min
Data Pitch	0.1 nm
Bandwidth	1 nm

6.4.1.1 Oligonucleotide Concentration Determination

A blank was recorded of 1000 µL of water. Samples recorded from solutions composed of 20 µL of DNA stock in 980 µL of water. The baseline was subtracted from the sample and zeroed at 400 nm. The absorbance at wavelength 260 nm was recorded and inserted into **Equation 6.2,** where A_{260} is the absorbance at 260 nm, *Df* is the dilution factor of 50, ε is the extinction coefficient of the oligonucleotide being analysed (**Table 6.6**), and l is the path length of 1 cm. The calculated concentration is in µM.

Table 6.6: *Table demonstrating the sequence and extinction coefficients of Oligonucleotides used throughout Chapter 3.*

Oligo Name	Sequence (5´ – 3´)	Extinction Coefficient (1/µM·cm)
COVID Trigger	AGG GTA AAC CAA ATA CC	178,300
SC1 Trigger	AGG GTT AAA CCA CCG CC	164,200
SC2 Trigger	AGG GTC CTT AAC TTG CC	156,300
Template COVID´-COVID´	GGT ATT TGG TTT ACC CTG TGA GAC TCT GGA ATT TGG TTT ACC CT	408,700
Template SC1´-SC1´	GGC GGT GGT TTA ACC CTG TGA GAC TCT GGC GGT GGT TTA ACC CT	407,200
Template SC2´-SC2´	GGC AAG TTA AGG ACC CTG TGA GAC TCT GGT ATT TGG TTT ACC CT	417,800
Binder DNA COVID	AGG GTA AAC CAA ATA CCT GGT GTA TAC GTT	303,600
Binder DNA SC1	AGG GTT AAA CCA CCG CCT GGA GAT CAA TTT	293,700
Binder DNA SC2	AGG GTC CTT AAC TTG CCT GGT TGT GAT GGT	280,200
Trigger N	AGG GTC TGA TTA GTT CC	164,100
Template N´-N´	GGA ACT AAT CAG ACC CTG TGA GAC TCT GGA ACT AAT CAG ACC CT	429,000
Binder DNA N	AGG GTC TGA TTA GTT CCT GGT CCC CAA AAT	287,100

6.4.2 PHE Samples

P2 stock of virus was acquired from High Containment Microbiology, PHE, Porton Down, corresponding to the SARS-CoV-2/human/AUS/VIC16832/2020 isolate, which was originally isolated in Australia from a COVID-19 patient in 2020. To prepare samples for RNA extraction, media containing the virus was added to Buffer AVL (Qiagen) in a 1/5 ratio and heated to 60 °C for 30 min in a calibrated heat block. Samples were then extracted on the MagNAPure96 (Roche) automated extraction system and then run on the Abbott M2000 RT-qPCR Test for SARS-CoV-2 RNA Detection. For EXPAR assay development, positive and negative RNA samples from the SARS-CoV-2 assays were separately combined in MagNA Pure elution buffer (giving 29,080 RNA copies per µL for the combined positive sample). Upon receipt from PHE, each sample (positive and negative) was diluted 400-fold with water, aliquoted into 50 µL vials, and stored at −80 °C.

6.4.3 Exponential Amplification Reaction

The protocol first involves the preparation of three solutions, Part A, Part B and Part C, followed by an addition step and then finally an amplification step.

6.4.3.1 Part A

1.50 µL of water, 2.50 µL of 10× Isothermal amplification buffer, 3.75 µL of BSA solution, 1.50 µL of *Bst* 2.0 DNA polymerase (1.6 U/µL) and then 0.75 µL of Nt.*Bst*NBI (10 U/µL).

6.4.3.2 Part B

6.30 µL of water, 5.00 µL of 10× Isothermal amplification buffer, 0.75 µL of **Template COVID’-COVID’** (1 µM), 2.40 µL of $MgSO_4$ (100 mM), 1.50 µL dNTP (10 nM), 0.75 µL of dsGreen (1:5 dilution in DMSO from 100× to 20×) and then 0.30 µL of SSB solution.

6.4.3.3 Part C

Sensitivity test (no RNA target): 3 µL of one trigger at **COVID Trigger** (100 nM, 10 nM, 1 nM, 100 pM, 10 pM, 1 pM, and a blank).

OR

Specificity test (no RNA target): 3 µL of one trigger at 100 nM (**COVID Trigger** or **SC1 Trigger** or **SC2 Trigger** or ***Chlamydia* Trigger**).

OR

RTF EXPAR assay (two-pot RTF-EXPAR): 10 µL of RNA:DNA heteroduplex digestion mixture, prepared as follows: 25 µL of water, 5 µL of 10× Isothermal amplification buffer, 5 µL *Bst*NI (10 U/µL), 10 µL of **Binder DNA COVID** (1 µM), and then 5 µL of positive or negative sample. The mixture is then incubated at 50 °C for 5 min.

OR

RTF EXPAR assay (one-pot RTF-EXPAR): Reagents are mixed together as follows: 1 µL *Bst*NI (2 U/µL), 2 µL of **Binder DNA COVID** (1 µM), and then 3 µL of positive or negative sample.

6.4.3.4 Addition Step

Part B (17 µL) is added to a PCR tube, and to this is added Part C (6 µL), followed by Part A (10 µL). The tube is then sealed, with the contents then subjected to amplification.

6.4.3.5 EXPAR Amplification

Isothermal incubation and fluorescence signal measurements are performed using an either an Agilent Mx3005P Real-Time PCR system or a Thermo Fisher QuantStudio 5 Real-Time PCR system, 96-well, 0.2 mL. The temperature is set at 25 °C for 15 seconds, before being raised to 50 °C for the duration of the assay, with the fluorescence reading measured every 10 seconds over an incubation time of 30 min.

6.4.4 Loop Mediated Isothermal Amplification

LAMP was performed according to the manufacturer's instructions. A reaction solution was composed in accordance with Table 6.7. 24 µL was transferred to a PCR tube sealed and vortexed. The reaction solution was incubated, and the fluorescence measured using a Thermo Fisher QuantStudio 5 Real-Time PCR system, 96-well, 0.2 mL. The temperature is set at 65 °C with the fluorescence reading measured every 10 seconds over an incubation time of 30 min.

Table 6.7: LAMP reaction mix.

Component	Volume Used
WarmStart LAMP 2x Master Mix	12.5 µL
Fluorescent Dye (50x)	0.5 µL
LAMP Primer Mix (10x)	2.5 µL
Target RNA	1 µL
Water	8.5 µL

6.4.5 Polymerase Chain Reaction

PCR was performed according to the manufacturer's instructions. 390 µL of rehydration buffer was added to the reaction mix tube. 15 µL of rehydrated reaction mix was added to a PCR tube, to this was added 5 µL of positive or negative sample before being sealed. The reaction solution was incubated according to Table 6.8, and the fluorescence measured using a Thermo Fisher QuantStudio 5 Real-Time PCR system, 96-well, 0.2 mL.

Table 6.8: PCR thermocycler conditions.

Temperature	Time	Cycles
45 °C	15 minutes	1
95 °C	2 minutes	1
95 °C	10 seconds	45
60 °C	50 seconds	

6.5 Chapter 5 Experimental

6.5.1 Foldamer Stock Preparation

0.03 g of foldamer P(*p*-AT) **S-1** was solvated in 1 mL of DMF to create a stock solution of concentration 923.1 µM.

6.5.2 Foldamer Sample Preparation

Samples were prepared by diluting the foldamer stock to 10 µM in DMF and water (**Table 6.9**). Water was added using a NE-1000 higher pressure programmable single syringe pump at a rate of 0.1 mL/min

Table 6.9*: Table showing the volume of stock solution, DMF and water used to create 1 mL aliquots of 10 μM P(p-AT) S-1 at varying water percentages.*

Percentage of Water	Volume of Stock solution (μL)	Volume of DMF (μL)	Volume of Water (μL)
0%	10.83	989.17	0
1%	10.83	979.17	10
2%	10.83	969.17	20
3%	10.83	959.17	30
4%	10.83	949.17	40
5%	10.83	939.17	50
6%	10.83	929.17	60
7%	10.83	919.17	70
8%	10.83	909.17	80
9%	10.83	899.17	90
10%	10.83	889.17	100
11%	10.83	879.17	110
12%	10.83	869.17	120
13%	10.83	859.17	130
14%	10.83	849.17	140
15%	10.83	839.17	150
16%	10.83	829.17	160
17%	10.83	819.17	170
18%	10.83	809.17	180
19%	10.83	799.17	190

20%	10.83	789.17	200
21%	10.83	779.17	210
22%	10.83	769.17	220
23%	10.83	759.17	230
24%	10.83	749.17	240
25%	10.83	739.17	250
26%	10.83	729.17	260
27%	10.83	719.17	270
28%	10.83	709.17	280
29%	10.83	699.17	290
30%	10.83	689.17	300
35%	10.83	639.17	350
40%	10.83	589.17	400
45%	10.83	539.17	450
50%	10.83	489.17	500
55%	10.83	439.17	550
60%	10.83	389.17	600
65%	10.83	339.17	650
70%	10.83	289.17	700
75%	10.83	239.17	750
80%	10.83	189.17	800
85%	10.83	139.18	850
90%	10.83	89.17	900

6.5.3 UV-Vis Spectroscopy

All measurements were performed on a Jasco V-750 UV-Visible spectrophotometer (Jasco, Japan) in a 1 mL quartz cuvette. Using the following parameters:

***Table 6.10**: Table of the parameters used during UV-vis measurements.*

Wavelength Range	270 – 380 nm
Scan Speed	50 nm/min
Data Pitch	0.1 nm
Bandwidth	1 nm

6.5.4 Circular Dichroism Spectroscopy

All measurements were performed on a Jasco J-1500 spectropolarimeter (Jasco, Japan) using 100 µL of sample in a 10 mm pathlength 6Q quartz cuvette. Baselines were recorded of the solvent system (DMF/H_2O) absent of analyte, the baseline was subtracted and zeroed at 380 nm. All samples were recorded using following parameters.

***Table 6.11:** Table of the parameters using during CD measurements.*

Wavelength Range	270 – 380 nm
Scan Speed	50 nm/min
Data Pitch	0.1 nm
Bandwidth	1 nm
Response Time	1 second
Accumulations	3
Scanning Mode	Continuous

6.5.5 Linear Dichroism Spectroscopy

All measurements were performed on a Jasco J-1500 spectropolarimeter (Jasco, Japan) which has been modified to measure LD.

6.5.5.1 Measurement of LD spectra using Couette Flow Alignment.

Samples (100 µL) were placed into a micro-Couette cell.[3] Baselines were recorded from a non-rotating cell; samples were then recorded from a rotating cell (3000 rpm) the baseline was subtracted and zeroed at 400 nm.

***Table 6.12**: Table of the parameters using during Couette flow LD measurements.*

Wavelength Range	270 – 380 nm
Scan Speed	50 nm/min
Data Pitch	0.1 nm
Bandwidth	1 nm
Response Time	1 second
Accumulations	3
Scanning Mode	Continuous

6.5.5.1.1 Shear Flow Linear Dichroism

Shear flow measurements were performed using the same conditions as Couette flow alignment, except spinning speed was varied from 0 – 5000 rpm at 1000 rpm intervals.

6.5.5.2 Polyethylene Oxidised Films

Polyethylene (PE) sheets with dimensions of 3 x 5 cm were placed in a Plasma Cleaner PDC-32G-2 connected to an oxygen gas supply for 4 minutes on high.

6.5.5.2.1 Measurement of LD spectra using Stretched Film Alignment

PE^{OX} sheets of dimensions 3 x 5 cm were placed between the jaws of a mechanical stretcher (**Figure 6.1**) built for this purpose at the University of Warwick. The film was them stretched from 2.5 cm to 4.5 cm. The analyte was added after the stretching. Baselines were collected on stretched films where a small quantity of solvent has been dropped on and allowed to dry. Samples (5 μL in solutions made of the baseline solvent) were dropped onto the stretched film, then allowed to dry. The stretched film and its holder were placed horizontally into the spectrometer so the horizontal light beam incident on the analyte was in the stretch direction. All samples were recorded using the following parameters.

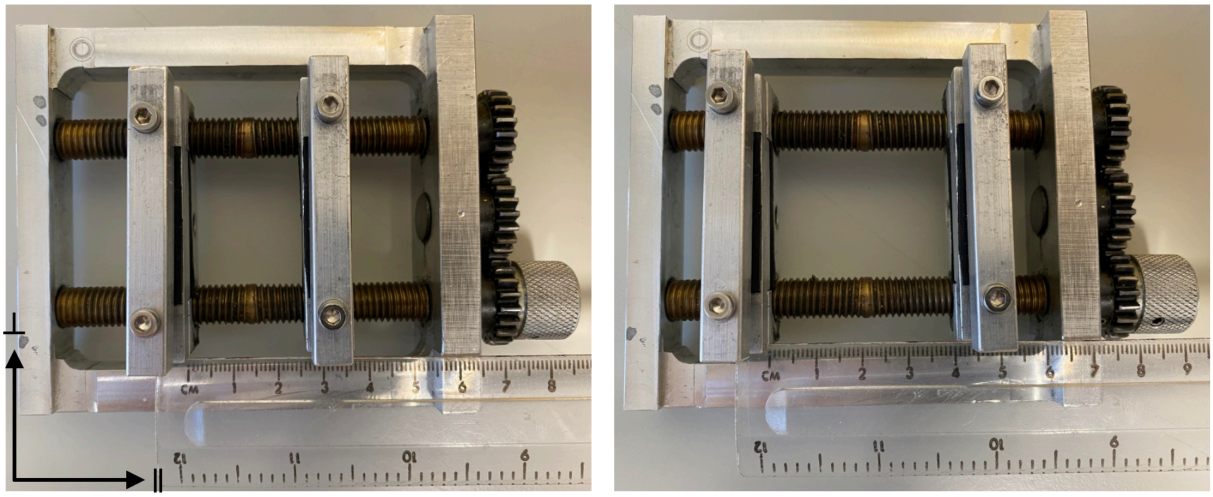


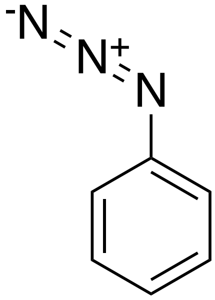
Figure 6.1: *(Left) Film stretcher prior to stretching. (Right) Film stretcher after stretching. Experimental orientation axes in relation to light polarisation: horizontal (stretch direction) is parallel (||) and vertical is perpendicular (⊥). Light is incident on the sample through the plane of the diagram.*

***Table 6.13**: Table of the parameters using during stretched film LD measurements.*

Wavelength Range	200 – 400 nm
Scan Speed	50 nm/min
Data Pitch	0.1 nm
Bandwidth	1 nm
Response Time	1 second
Accumulations	3
Scanning Mode	Continuous

6.5.6 Synthesis

6.5.6.1 Phenyl Azide, 5.2

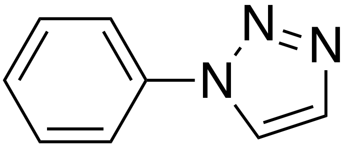


A known compound synthesised according to a literature procedure.[4]

To a solution of aniline (**5.1)** (2 g, 21.48 mmol) in H_2O (30 mL) was added concentrated HCl (9 mL), $NaNO_2$ (1.78 g, 25.80 mmol) at 0 °C. This reaction was stirred at 0 °C for 0.5 h. A solution of NaN_3 (2.37 g, 36.46 mmol) in H_2O (10 mL) was added dropwise to the mixture at 0 °C for 0.5 h and allowed to warm to room temperature overnight. DCM (50 mL) was added, the organic phase separated and the aqueous washed with DCM (2 x 30 mL). The combined organic layers were dried over $MgSO_4$, filtered, and concentrated to yield **5.2** as an orange liquid (2.40 g, 94%). **FTIR** ν_{max} (neat)/cm^{-1}: 2123 (s), 1593 (s), 1293 (s), 1279 (s), 744 (s); **^{1}H NMR** (400 MHz, CDCl3): δ (ppm) = 7.31 – 7.22 (m, 2H), 7.10 – 7.02 (m, 1H), 6.97 – 6.91 (m, 2H); **^{13}C NMR** (101 MHz, CDCl3): δ (ppm) = 140.10, 129.85, 124.96, 119.12; **HRMS** (ESI): m/z calculated for $C_6H_5N_3$: 119.05 [M]; found: 119.01 [M].

Analytical data in agreement with literature values.[4]

6.5.6.2 1-Phenyl-1,3,3-triazole, PT

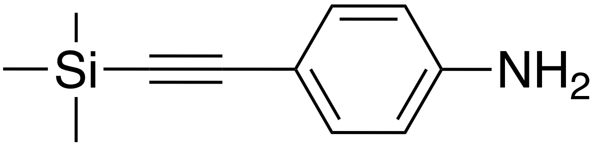


A known compound synthesised according to a literature procedure.[5]

A flask was charged with a MeCN (80 mL) and H_2O (40 mL), **5.2** (0.36 g, 3.00 mmol), copper iodide (190 mg, 1.0 mmol), sodium ascorbate (198 mg, 1.0 mmol) and calcium carbide (0.54 g, 8.5 mmol). The mixture was stirred at room temperature overnight. After the reaction was completed, the system was acidified to pH 5.0 with HCl (2 M). The mixture was extracted with ethyl acetate (3 x 30 mL). The organic layer was separated, washed with water (50 mL) and brine (50 mL), dried over $MgSO_4$, and filtered. Evaporation of the solvent yielded the crude product, which was subjected to flash column chromatography (SiO_2, Diethyl Ether/Hexane = 1:1) to afford **PT** as an orange liquid (0.18 g, 41%). **FTIR** ν_{max} (neat)/cm^{-1}: 3022 (w), 2253 (s), 1443 (br), 758 (s); **^{1}H NMR** (400 MHz, $CDCl_3$): δ (ppm) = 8.00 (d, J = 1.1 Hz, 1H), 7.85 (d, J = 1.1 Hz, 1H), 7.77 – 7.73 (m, 2H), 7.57 – 7.51 (m, 2H), 7.48 – 7.42 (m, 1H); **^{13}C NMR** (101 MHz, $CDCl_3$): δ (ppm) = 134.69, 129.98, 128.99, 121.91, 120.89; **HRMS** (ESI): *m/z* calculated for $C_8H_7N_3$: 145.06 [M]; found: 145.08 [M]

Analytical data in agreement with literature values.[5]

6.5.6.3 4-((trimethylsilyl)ethynyl)aniline, 5.5

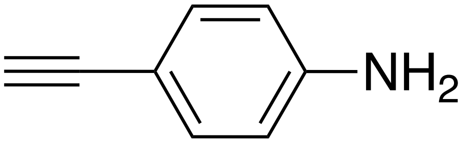


A known compound was synthesised according to a literature procedure.[6]

To a mixture of 4-iodoaniline (**5.4)** (2.19 g, 10 mmol), bis(triphenylphosphine)palladium chloride (70 mg, 0.10 mmol), copper iodide (20 mg, 0.10 mmol), triethylamine (11 mL) and THF (11 mL) under argon, trimethylsilyl acetylene (2.14 mL, 15 mmol) was added. The mixture was stirred at 45 °C for 24 hours. After cooling to room temperature, the mixture was filtered, extracted with ethyl acetate (2 x 25 mL), washed with aqueous 2 M HCl (25 mL) and brine (25 mL). The organic fraction was dried over $MgSO_4$, filtered, and concentrated *in vacuo*. After the removal of the solvent the residue was purified by column chromatography (SiO_2, Ethyl Acetate/Hexane = 1:3) to yield the **5.5** as a yellow solid. (1.67 g, 88%). **m.p.** 92 – 96 °C; **^{1}H NMR** (400 MHz, $CDCl_3$): δ (ppm) = 7.29 – 7.21 (m, 2H), 6.57 – 6.53 (m, 2H), 3.77 (s, 2H), 0.20 (s, 9H); **^{13}C NMR** (101 MHz, $CDCl_3$): δ (ppm) = 146.98, 133.59, 114.73, 112.70, 106.19, 91.58, 0.36; **HRMS** (ESI): *m/z* calculated for $C_{11}H_{15}NSi$: 189.33 [M]; found: 190.11 [M + H]

Analytical data in agreement with literature values.[6]

6.5.6.4 4-ethynylaniline, 5.6

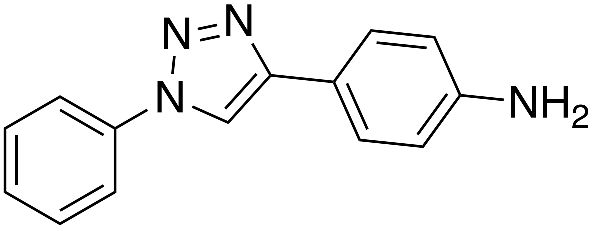


A known compound was synthesised according to a literature procedure.[6]

To a solution of **5.5** (1.1 g, 5.8 mmol), in methanol (46 mL) was added potassium carbonate (0.1 g, 0.74 mmol) and stirred for 2 hours at room temperature under N_2. The reaction mixture was concentrated *in vacuo*, diluted with diethyl ether (50 mL), washed with water (2 x 25 mL), brine (25 mL). The organic layer was dried over $MgSO_4$, filtered, and concentrated to yield **5.6** as a beige solid. (0.54 g 79%) **m.p.** = 106 – 110 °C; **^{1}H NMR** (400 MHz, $CDCl_3$): δ (ppm) = 7.33 – 7.27 (m, 2H), 6.62 – 6.56 (m, 2H), 3.81 (s, 2H), 2.96 (s, 1H); **^{13}C NMR** (101 MHz, $CDCl_3$): δ (ppm) = 147.21, 133.66, 114.77, 111.50, 84.59, 75.11; HRMS (ESI): *m/z* calculated for C_8H_7N: 117.15 [M]; found: 117.05 [M].

Analytical data in agreement with literature values.[6]

6.5.6.5 4-(1-phenyl-1*H*-1,2,3-triazol-4-yl)aniline, PTA



A known compound was synthesised according to a literature procedure.[7]

A round bottomed flask was charged with a mixture of *tert*-butyl alcohol (2.5 mL) and water (2.5 mL). Then **5.2** (0.24 g, 2 mmol), **5.6** (0.23 g, 2 mmol), 1 M sodium ascorbate solution (0.2 mL) and $CuSO_4.5H_2O$ (5 mg, 0.02 mmol) was added. The suspension was stirred at 60 °C for 24 hours, after cooling to room temperature 5 mL of water was added. The mixture was filtered and washed with water to yield **PTA** as a beige solid. Liquid-liquid diffusion between water and DMSO produced single yellow crystals which were suitable for X-ray analysis. (0.80 g 95.0%) **m.p.** = 186 – 189 °C; **^{1}H NMR** (400 MHz, DMSO-*d$_6$*): δ (ppm) = 8.99 (s, 1H), 7.97 – 7.88 (m, 2H), 7.61 (t, *J* = 8.3 Hz, 4H), 7.48 (tt, *J* = 6.9, 1.1 Hz, 1H), 6.65 (d, *J* = 8.5 Hz, 2H), 5.33 (s, 2H); **^{13}C NMR** (101 MHz, DMSO-*d$_6$*): δ (ppm) = 148.98, 148.40, 136.84, 129.96, 128.49, 126.46, 119.82, 117.78, 117.21, 114.04; **HRMS** (ESI): *m/z* calcd for $C_{14}H_{12}N_4$: 236.28 [M]; found: 237.11 [M].

Analytical data in agreement with literature values.[7]

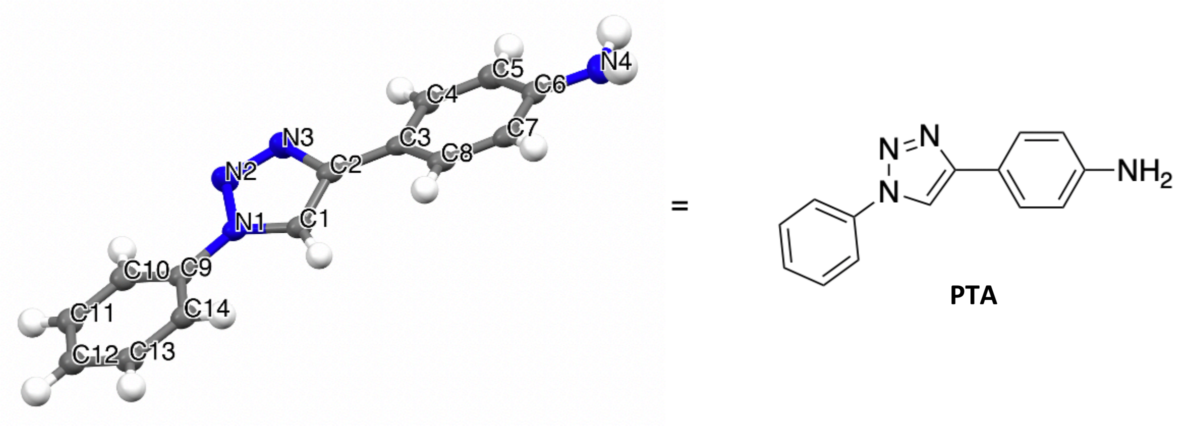
6.6 References

1. Little, H. A. The Development of Novel Diagnostic Sensors Based on Linear Dichroism Spectroscopy. (University of Birmingham, 2016).
2. Leonard, N. M. & Brunckova, J. In Situ Formation of N-Trifluoroacetoxy Succinimide (TFA-NHS): One-Pot Formation of Succinimidyl Esters, N-Trifluoroacetyl Amino Acid Succinimidyl Esters, and N-Maleoyl Amino Acid Succinimidyl Esters. *The Journal of Organic Chemistry* **76**, 9169–9174 (2011).
3. Marrington, R. *et al.* Validation of new microvolume Couette flow linear dichroism cells. *The Analyst* **130**, 1608 (2005).
4. Qin, Y. *et al.* Design, synthesis and antibacterial evaluation of novel 15-membered 11a-azahomoclarithromycin derivatives with the 1, 2, 3-triazole side chain. *European Journal of Medicinal Chemistry* **180**, 321–339 (2019).
5. Jiang, Y., Kuang, C. & Yang, Q. The Use of Calcium Carbide in the Synthesis of 1-Monosubstituted Aryl 1,2,3-Triazole via Click Chemistry. *Synlett* **2009**, 3163–3166 (2009).
6. Arakawa, Y., Kang, S., Tsuji, H., Watanabe, J. & Konishi, G. The design of liquid crystalline bistolane-based materials with extremely high birefringence. *RSC Adv.* **6**, 92845–92851 (2016).
7. Henkes, L. M., Haus, P., Jäger, F., Ludwig, J. & Meyer-Almes, F.-J. Synthesis and biochemical analysis of 2,2,3,3,4,4,5,5,6,6,7,7-dodecafluoro-N-hydroxy-octanediamides as inhibitors of human histone deacetylases. *Bioorganic & Medicinal Chemistry* **20**, 985–995 (2012).

Chapter 7: Appendices

7.1 Chapter 5 Appendices

7.1.1 X-Ray Data for PTA



PTA

Table 7.1: Crystal data and structure refinement for PTA.

Identification code	JC228	
Empirical formula	$C_{14}H_{12}N_4$	
Formula weight	236.28	
Temperature	100.01(10) K	
Crystal system	Monoclinic	
Space group	$P2_1/n$	
Unit cell dimensions	a = 5.7333(3) Å	α = 90°
	b = 10.4071(7) Å	β = 91.142(6)°
	c = 19.1737(13) Å	γ = 90°
Volume	1143.81(12) Å^3	
Z	4	
Density (calculated)	1.372 g/cm^3	
Absorbance coefficient	0.685 mm^{-1}	
F(000)	496.0	
Crystal size	0.234 × 0.054 × 0.029 mm^3	
Radiation	Cu Kα (λ = 1.54184)	

2Θ range for data collection	9.226 to 145.094°
Index ranges	-4 ≤ h ≤ 6, -12 ≤ k ≤ 12, -21 ≤ l ≤ 23
Reflections collected	4167
Independent reflections	2199 [R_{int} = 0.0262, R_{sigma} = 0.0294]
Data/restraints/parameters	2199/0/171
Goodness-of-fit on F^2	1.121
Final R indexes [I>=2σ (I)]	R_1 = 0.0521, wR_2 = 0.1457
Final R indexes [all data]	R_1 = 0.0588, wR_2 = 0.1509
Largest diff. peak/hole	0.34/-0.30 e·Å^{-3}

***Table 7.2**: Fractional Atomic Coordinates (×10^4) and Equivalent Isotropic Displacement Parameters (Å^2×10^3) for JC228. U_{eq} is defined as 1/3 of the trace of the orthogonalised U_{IJ} tensor.*

	x	*y*	*z*	**U(eq)**
C(1)	5359(3)	3448.6(19)	4508.7(10)	24.0(4)
C(2)	4090(3)	4174(2)	4037.0(11)	24.6(4)
C(3)	4849(3)	5180.0(19)	3560.2(11)	24.2(4)
C(4)	3436(3)	5554(2)	2987.6(11)	26.2(5)
C(5)	4141(4)	6492(2)	2531.9(11)	29.0(5)
C(6)	6296(3)	7109(2)	2620.6(11)	26.1(5)
C(7)	7718(3)	6731(2)	3190.4(11)	26.8(5)
C(8)	7005(3)	5794(2)	3645.3(11)	25.6(4)
C(9)	4121(3)	1707.8(19)	5337.4(10)	23.3(4)
C(10)	2357(4)	831(2)	5456.2(11)	27.3(5)
C(11)	2632(4)	-81(2)	5971.9(11)	29.6(5)
C(12)	4673(4)	-127(2)	6376.3(12)	29.7(5)

C(13)	6431(4)	759(2)	6249.2(11)	29.3(5)
C(14)	6185(3)	1680(2)	5733.4(11)	26.5(5)
N(1)	3795(3)	2661.5(16)	4811.8(9)	23.3(4)
N(2)	1630(3)	2878.9(19)	4539.0(10)	32.0(4)
N(3)	1819(3)	3784.3(19)	4073.0(10)	31.6(4)
N(4)	7028(4)	8008(2)	2149.4(11)	33.8(5)

Table7.3: *Anisotropic Displacement Parameters* ($Å^2\times10^3$) *for JC228. The Anisotropic displacement factor exponent takes the form:* $-2\pi^2[h^2a^{*2}U_{11}+2hka^*b^*U_{12}+\dots]$

	U_{11}	**U_{22}**	**U_{33}**	**U_{23}**	**U_{13}**	**U_{12}**
C(1)	18.4(9)	24.2(10)	29.4(10)	-2.4(8)	-0.2(7)	1.1(7)
C(2)	21.3(10)	23.5(10)	28.9(10)	-4.5(8)	-0.5(7)	3.4(8)
C(3)	19.6(10)	23.8(10)	29.0(10)	-4.9(8)	-0.1(7)	5.2(7)
C(4)	19.6(10)	27.7(10)	31.2(11)	-1.9(8)	-0.9(7)	-0.1(8)
C(5)	24.5(10)	31.7(11)	30.6(10)	-3.5(9)	-3.4(8)	1.8(8)
C(6)	25.0(10)	23.9(10)	29.7(10)	-2.6(8)	4.6(8)	1.7(8)
C(7)	18.7(10)	25.9(10)	35.8(11)	-6.7(8)	1.2(8)	1.4(8)
C(8)	21.9(10)	24.3(10)	30.6(10)	-4.7(8)	-2.0(7)	4.5(8)
C(9)	20.5(9)	23.6(10)	25.8(10)	-3.1(8)	2.3(7)	2.8(7)
C(10)	22.0(10)	27.1(11)	32.8(11)	-5.1(8)	-1.4(8)	-1.6(8)
C(11)	26.4(11)	27.2(11)	35.1(12)	-0.6(9)	1.9(8)	-1.7(8)
C(12)	27.3(11)	29.5(11)	32.4(11)	1.6(9)	2.1(8)	3.3(8)
C(13)	22.2(10)	33.3(11)	32.3(11)	-0.9(9)	-3.0(8)	2.9(8)
C(14)	19.1(9)	28.0(11)	32.4(11)	-3.8(8)	1.7(7)	-0.4(8)

N(1)	17.5(8)	24.3(8)	28.2(9)	-2.9(7)	0.0(6)	0.9(6)
N(2)	20.7(9)	34.7(10)	40.4(10)	9.9(8)	-3.6(7)	0.5(7)
N(3)	20.1(9)	34.2(10)	40.4(10)	8.9(8)	-2.2(7)	-0.9(7)
N(4)	32.5(10)	32.1(10)	37.0(11)	3.3(8)	2.8(8)	-4.2(8)

Table 7.4: *Bond Lengths (Å) for JC228.*

C(1) – C(2)	1.375(3)	C(7) – C(8)	1.376(3)
C(1) – N(1)	1.354(3)	C(9) – C(10)	1.385(3)
C(2) – C(3)	1.462(3)	C(9) – C(14)	1.393(3)
C(2) – N(3)	1.367(3)	C(9) – N(1)	1.424(3)
C(3) – C(4)	1.407(3)	C(10) – C(11)	1.378(3)
C(3) – C(8)	1.398(3)	C(11) – C(12)	1.392(3)
C(4) – C(5)	1.376(3)	C(12) – C(13)	1.391(3)
C(5) – C(6)	1.400(3)	C(13) – C(14)	1.382(3)
C(6) – C(7)	1.406(3)	N(1) – N(2)	1.356(2)
C(6) – N(4)	1.372(3)	N(2) – N(3)	1.304(3)

Table 7.5: *Bond Angles (°) for JC228*

N(1) – C(1) – C(2)	105.54(17)	C(10) – C(9) – C(14)	120.8(2)
C(1) – C(2) – C(3)	130.09(19)	C(10) – C(9) – N(1)	119.31(18)
N(3) – C(2) – C(1)	107.23(18)	C(14) – C(9) – N(1)	119.89(18)
N(3) – C(2) – C(3)	122.68(18)	C(11) – C(10) – C(9)	119.93(19)
C(4) – C(3) – C(2)	120.82(18)	C(10) – C(11) – C(12)	120.4(2)

C(8) – C(3) – C(2)	121.86(18)	C(13) – C(12) – C(11)	119.0(2)
C(8) – C(3) – C(4)	117.32(19)	C(14) – C(13) – C(12)	121.40(19)
C(5) – C(4) – C(3)	121.33(19)	C(13) – C(14) – C(9)	118.52(19)
C(4) – C(5) – C(6)	121.2(2)	C(1) – N(1) – C(9)	130.27(16)
C(5) – C(6) – C(7)	117.6(2)	C(1) – N(1) – N(2)	110.02(17)
N(4) – C(6) – C(5)	120.9(2)	N(2) – N(1) – C(9)	119.71(17)
N(4) – C(6) – C(7)	121.49(19)	N(3) – N(2) – N(1)	107.27(17)
C(8) – C(7) – C(6)	121.06(19)	N(2) – N(3) – C(2)	109.94(17)
C(7) – C(8) – C(3)	121.54(19)		

Table 7.6: *Torsion Angles (°) for JC228.*

C(1) – C(2) – C(3) – C(4)	161.9(2)	C(10) – C(9) – C(14) – C(13)	-0.2(3)
C(1) – C(2) – C(3) – C(8)	-17.1(3)	C(10) – C(9) – N(1) – C(1)	-165.8(2)
C(1) – C(2) – N(3) – N(2)	0.5(2)	C(10) – C(9) – N(1) – N(2)	13.9(3)
C(1) – N(1) – N(2) – N(3)	0.0(2)	C(10) – C(11) – C(12) – C(13)	-0.3(3)
C(2) – C(1) – N(1) – C(9)	-179.96(19)	C(11) – C(12) – C(13) – C(14)	0.2(3)
C(2) – C(1) – N(1) – N(2)	0.3(2)	C(12) – C(13) – C(14) – C(9)	0.0(3)
C(2) – C(3) – C(4) – C(5)	-179.43(19)	C(14) – C(9) – C(10) – C(11)	0.1(3)
C(2) – C(3) – C(8) – C(7)	179.34(19)	C(14) – C(9) – N(1) – C(1)	15.3(3)
C(3) – C(2) – N(3) – N(2)	179.99(18)	C(14) – C(9) – N(1) – N(2)	-164.98(19)
C(3) – C(4) – C(5) – C(6)	0.1(3)	N(1) – C(1) – C(2) – C(3)	-179.91(19)
C(4) – C(3) – C(8) – C(7)	0.3(3)	N(1) – C(1) – C(2) – N(3)	-0.5(2)
C(4) – C(5) – C(6) – C(7)	0.3(3)	N(1) – C(9) – C(10) – C(11)	-178.77(19)
C(4) – C(5) – C(6) – N(4)	177.5(2)	N(1) – C(9) – C(14) – C(13)	178.70(18)

C(5) – C(6) – C(7) – C(8)	-0.4(3)	N(1) – N(2) – N(3) – C(2)	-0.3(2)
C(6) – C(7) – C(8) – C(3)	0.1(3)	N(3) – C(2) – C(3) – C(4)	-17.4(3)
C(8) – C(3) – C(4) – C(5)	-0.4(3)	N(3) – C(2) – C(3) – C(8)	163.5(2)
C(9) – C(10) – C(11) – C(12)	0.1(3)	N(4) – C(6) – C(7) – C(8)	-177.53(19)
C(9) – N(1) – N(2) – N(3)	-179.78(17)		

Table 7.7: *Hydrogen Atom Coordinates (Å×10⁴) and Isotropic Displacement Parameters (Å²×10³) for JC228.*

	x	***y***	***z***	**U(eq)**
H(1)	6956.38	3490.03	4600.32	29
H(4)	1995.35	5160.27	2915.53	31
H(5)	3168.45	6720.59	2158.23	35
H(7)	9163.9	7120	3260.81	32
H(8)	7979.89	5562.99	4018.02	31
H(10)	988.21	857.78	5188.13	33
H(11)	1446.11	-670.74	6050.17	35
H(12)	4858.69	-739.3	6726.1	36
H(13)	7800.55	731.32	6516.94	35
H(14)	7371.03	2268.1	5652.51	32
H(4A)	5890(60)	8360(30)	1817(16)	50(8)
H(4B)	8150(60)	8590(30)	2303(16)	51(8)

7.1.2 NMR Data for stretched Film LD monomers

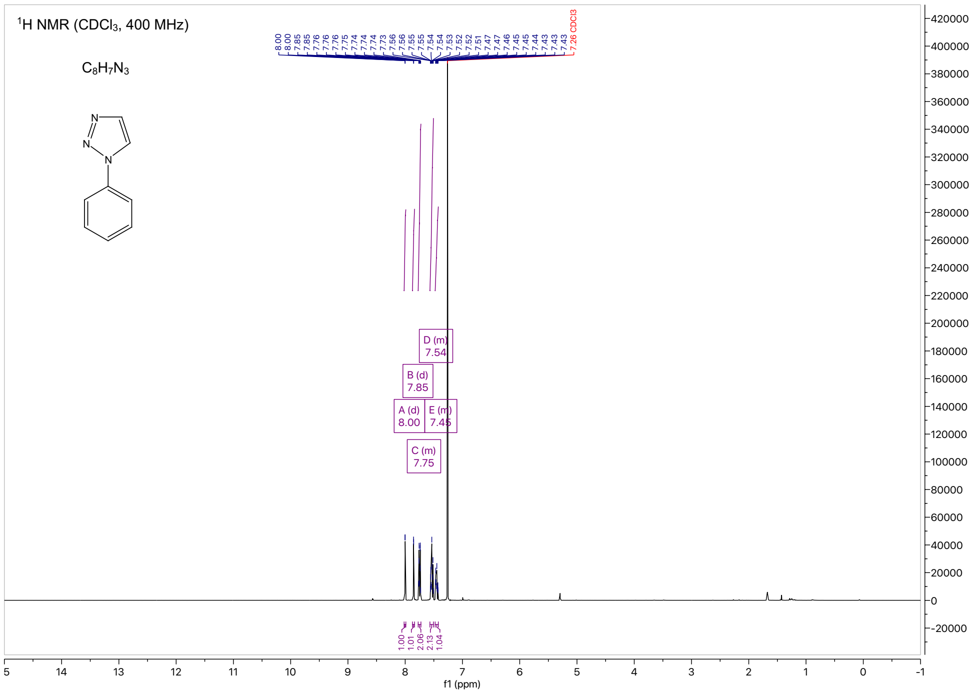


Figure 7.1: *¹H NMR for **PT** in CDCl₃*

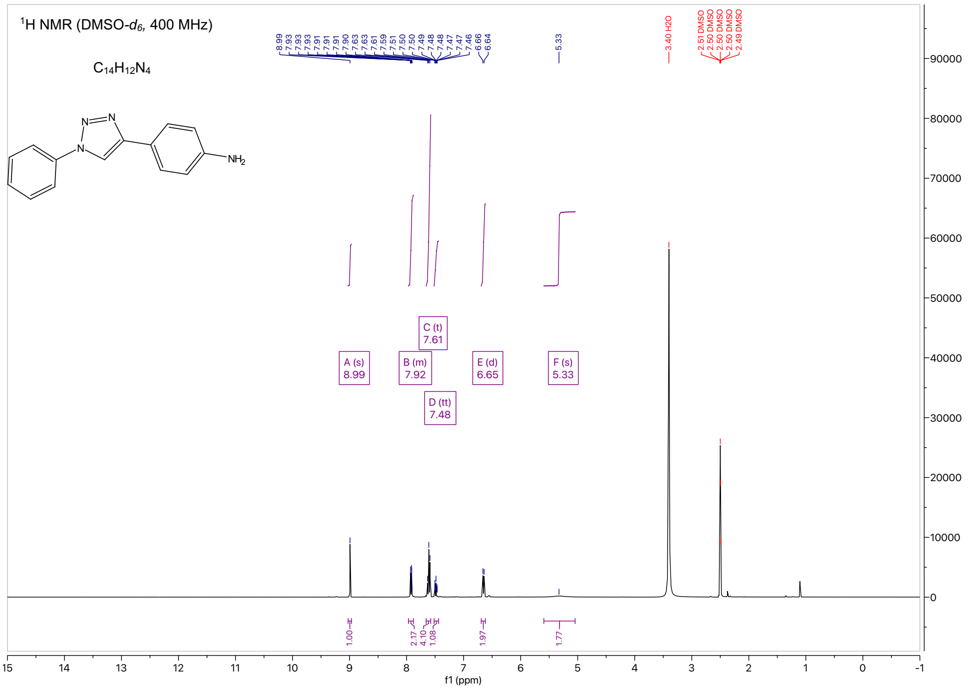


Figure 7.2: *¹H NMR for **PTA** in DMSO*