Designer DNA for Sensing and Nanomaterials

Rachel Clare Little

A thesis submitted for the degree of

Doctor of Philosophy in Chemistry

School of Natural and Environmental Science

Newcastle University

June 2019

Abstract

The Whitfield enzymatic extension method was explored for the development of three key concepts; i) amplification of size specific designer DNA (improved yield), ii) the synthesis of known spatial binding sites (novel nanomaterials) and iii) extended DNA brush synthesis (sensing applications).

The size recovery and amplification of specific lengths of designer DNA products were investigated for use in DNA nanomaterials. Initially, oligoseeds, $[AT]_{10}/[TA]_{10}$ and $[GC]_{10}/[GC]_{10}$, were enzymatically extended using the Whitfield method to give large distributions of DNA of up to 20,000 bp. Product concentrations were 71.4 ng/μL for $[AT]_n$ and 135 ng/μL for $[GC]_n$. Size recovery by removing aliquots of size specific DNA during gel electrophoresis afforded DNA aliquots of narrow size and in low concentration, < 5 ng/μL, thus amplification by ligation and qPCR was used to increase the concentration of the final product to ~ 30 ng/μL.

In addition, the synthesis of long DNA products containing 3 multi-base sequences, ABC was developed. This approach afforded the incorporation of a 4^{th} sequence, D. The ABCD-DNA method produced various sequences and lengths of novel designer DNA ranging from 500 - 5,000 bp. The new sequence D, not in the original oligoseed, allowed for the successful inclusion of spatially separated sequences of phosphorothioate bases into ssDNA or dsDNA of 500 - 1,500 bp. 3 nm gold nanoparticles bound specifically to the phosphorothioate regions as seen by AFM.

Enzymatic extension of surface immobilised oligoseeds was used to grow long repeat sequence DNA brushes for enhanced sensing. The initial 20 bp oligoseed surface gave an enhanced fluorescent signal from the PicoGreen fluorescent indicator due to the multiplexed repeat sequences upon extension. The surface extension of the micro-satellite BAT-25 sequence demonstrated that small numbers of base mis-matches could be distinguished by comparison of their hybridised fluorescent signals.

Acknowledgements

I would like to thank Dr. Andrew Pike for providing me with the opportunity to undertake my PhD within his group and for all his support and supervision throughout my time at Newcastle. I would also like to thank Dr. Eimer Tuite who always offered valuable guidance which steered me through the more difficult times during my project. Furthermore, I am grateful to the School of Natural and Environmental Science (Chemistry) for funding my research.

I am eternally appreciative of the Pike group, Dr. Colette Whitfield, Dr. Samantha Lunn and Dr. Tom Bamford for their never-ending support and friendship, during and beyond my studies. My PhD experience was greatly enhanced by the friends I made on the Chem Nano corridor, many of whom I could not have completed my research without.

I would also like to extend my thanks to Prof Ijiro and his group at Hokkaido University for allowing me to work in their laboratory for 2 months during my studies.

Finally, I would not have been able to complete my research without the support of my family and my fiancé Stephen who were with me every step of the way.

Contents

Abbreviations

1D	One dimensional
2D	Two dimensional
3D	Three dimensional
5-Br-dU	5-bromo-deoxyuracil
5-C8-alkyne-dC	5-octyne-deoxycytidine
5-I-dC	5-iodo-deoxycytidine
6-S-dG	6-thiol-deoxyguanine
7-deaza-7-I-dA	7-deaza-7-iodo-deoxyadenine
A	Adenine
AFM	Atomic force microscopy
Ag(I)	Silver(I)
Am	Amine
APEGDMES	Acetalpolyethyleneglycolmethylethoxysilane
Au	Gold
AuNP	Gold nanoparticle
BAP	Bacterial alkaline phosphate
BSA	Bovine serum albumin
bT	Bicyclic thymine
C	Cytosine/Carbon
CF	Cystic fibrosis
CFTR	Cystic fibrosis transmembrane regulator
CMETS	2-(carbomethoxy)ethyltrichlorosilane
CPG	Controlled pore glass
dA	Deoxyriboadenine
dbDNA	Doggybone DNA™
dC	Deoxyribocytosine
dG	Deoxyriboguanosine
DNA	Deoxyribonucleic acid
dNTP	Deoxynucelotide triphosphate
dsDNA	Double stranded DNA
dT	Deoxyribothymine
EDC	*N*-(3-dimethylaminopropyl)-*N′*-ethylcarbodiimide hydrochloride

EDTA	Ethylenediaminetetraacetic acid
exo-	Exonuclease minus
extDNA	Extended DNA
FBA	Fluorescent base analogue
FcDNA	Ferrocene nucleic acid
FTIR	Fourier transform infra-red
G	Guanine
G4	G-quadruplex
H	Hydrogen
HPLC	High performance liquid chromatography
Im	Imidazole
IPA	Isopropanol
iSpPC	Photocleavable Linker
MD	Molecular dynamic
MMR	Mis-match repair
MSI	Microsatellite instability
N	Nitrogen
O	Oxygen
oligos	Oligonucleotides
P	Pyridine
pA	Pentacyclic adenine
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PEG	Polyethyleneglycol
PG	Pico Green
PME	Particle Mesh-Ewald
PNA	Polyamide nucleic acid
Pyr	Pyramidine
qPCR	Real-time PCR
RCA	Rolling circle amplification
S	Thiol/Sulfur
SAv	Streptavidin
SDS	Sodium Dodecyl Sulfate
Si	Silicon

ss-extDNA	Single stranded extended DNA
ssDNA	Single stranded DNA
T	Thymine
TBE	Tris, boric acid and EDTA
TEM	Tunnelling electron microscopy
Tm	Melting temperature
TP	Triphosphate
Tri	Triazole
UV-Vis	Ultraviolet-visible
αS	Phosphorothioate

Chapter 1

Introduction

Contents

1.1 Introduction

Beyond holding genetic information, DNA has been utilised in nanomaterial synthesis and sensing applications, e.g. one dimensional (1D),[1–3] two dimensional (2D)[4–6] and three dimensional (3D) nanomaterials,[7–9] DNA origami,[5] conductive nanowires[10–12] and diagnostic sensing.[13] DNA is integral to the fabrication of such nanomaterials, by imparting unique characteristics such as ease of synthesis, structural stability, molecular recognition, programmability and directed self-assembly.[14–16] Furthermore, using DNA as a material is advantageous due to its ability to assemble predictable nanoscale structures through the hydrogen bonding that arises from the complementary base pairing. [16,17]

1D nanomaterials are generally created by DNA longer than that of the persistence length (50 nm),[18] *e.g.* Lambda DNA, so they can be easily visualised. The majority of cases use commercially available DNA to produce 1D nanowires, where the conductive[19] or insulating[20] properties are dependent on the individual DNA components.[21] The Porath group reported the conductive nature of G-quadruplex DNA[22] as it provided a rigid structure with few defects and an additional route for charge transport to occur.[23]

DNA has also been used as a template for the deposition of a conductive material, most commonly metal ions.[20] The most popular metals for DNA templating which encourage long range electronic transport are silver,[20] palladium,[24] copper[25] and platinum.[26] During nanowire fabrication, the metal cation electrostatically associates with the phosphate backbone, and the subsequent reduction of the metal cation leaves a conductive metallic coating templated on the DNA, **Figure 1.1**. The conductive ability of the 1D nanowires may then be measured.

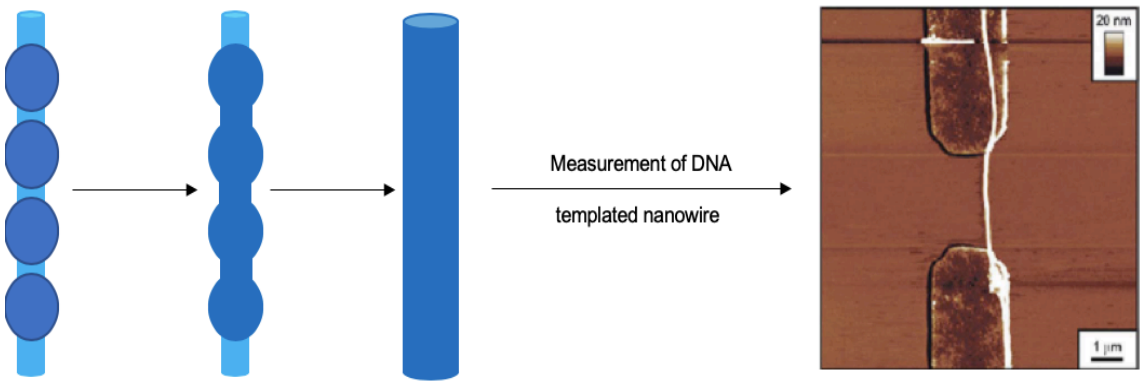


Figure 1.1. *Metal templating of a 1D nanowire to improve conductivity.*[1,27]

The specific base pairing of the complementary bases; adenine-thymine (A-T) and guanine-cytosine (G-C), leads to DNA self-assembly of synthetically made short oligonucleotides (oligos) to generate differing geometries. Seeman *et al.*[2,28] reported that DNA self-assembly could be used to fabricate thermodynamically stable nanostructures through branched junctions, **Figure 1.2**. To create self-assembled structures and devices, these branched junctions were coupled with sequence-specific double stranded DNA (dsDNA) strands which contained a region of protruding unpaired nucleotides on one of the strands creating an overhang, defined as 'sticky ends'.[28]

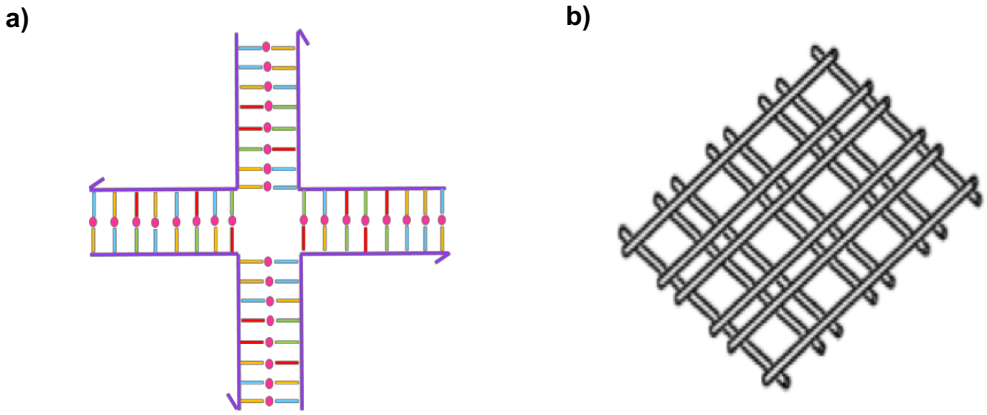
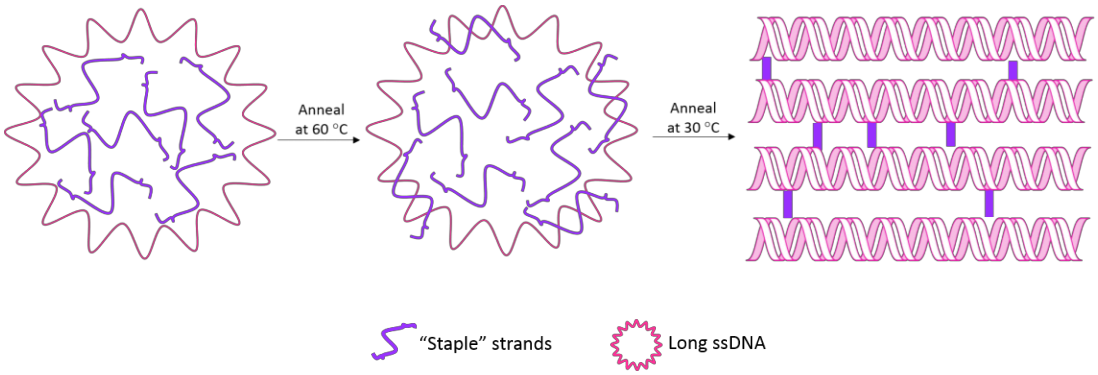


Figure 1.2. a) *DNA branched junction compiled of four DNA strands. Half-arrow indicates the 3' end of the strand and hydrogen bonding between the base pairs forming the double helix is shown by the pink dot.*[28] ***b)*** *DNA branched junctions formed into a tile structure.*[29]

Seeman *et al.*[16] advanced the idea of using self-assembly to create tile structures comprised of 'sticky ended' DNA, which hybridised to link the double helices together to produce cube-like structures. The concept of tile-based motifs was progressed, and studies were conducted leading to the introduction of DNA lattices which could be constructed into large 1D and 2D DNA arrays.[29]

In 2006, DNA origami was introduced by Rothemund *et al.*[5,30] and defined as when a long single stranded DNA (ssDNA) strand is used as a "scaffold", before being folded into a defined shape by the use of shorter DNA "staple" strands. These "staple" strands are complementary to at least two regions of the longer ssDNA, enabling hydrogen bonding to occur. DNA origami is assembled when the long ssDNA is placed in excess of the "staple" strands and annealed by heating in a buffer with a high concentration of magnesium, **Scheme 1.1**.[5,15]



Scheme 1.1. *DNA origami assembly. Initial annealing at 60 °C where "staple" stands attach onto the long strand before slow cooling to 30 °C where the origami structures stabilises.*[5,15]

DNA origami is used to assemble 2D and 3D nanostructures. Rothemund *et al.*[5] demonstrated the versatility of the method by constructing different 2D shapes such as squares, rectangles, smiley faces, stars and other complex structures, **Figure 1.3**.

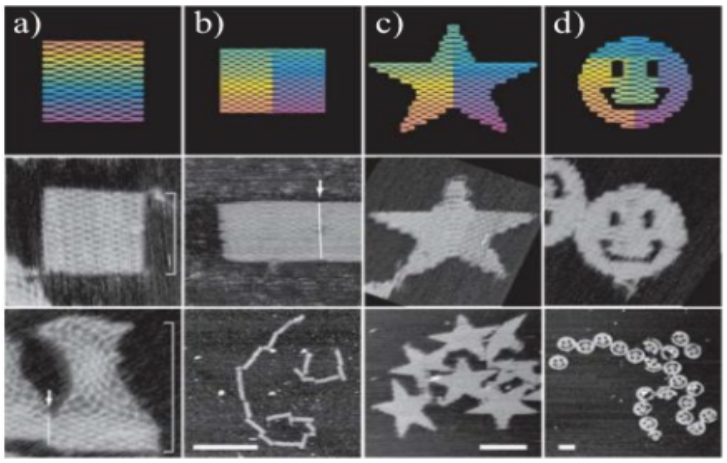


Figure 1.3. *DNA origami shapes. The top row shows the folding patterns and the bottom two rows show single and multiple origami shapes.* ***a)*** *square* ***b)*** *rectangle* ***c)*** *star and* ***d)*** *smiley face.*[5]

3D nanomaterials constructed by DNA origami can be produced by several methods, all of which require both long and short DNA lengths. The method proposed by Rothemund *et al.*[5] was to use a flat multi-domain origami to fold flat surfaces back to back through the stacking of the helices in separate domains. This method has been used to produce multiple styles of a 3D DNA origami box, **Figure 1.4**.[9,31] Another suggested method to produce 3D structures was to use multiple layers of DNA origami.[32] However, although the structures produced by multilayer origami were more rigid and stable, longer annealing times were required to form the final shape which was suggested to be due to obstructed diffusion inside the shapes.[15]

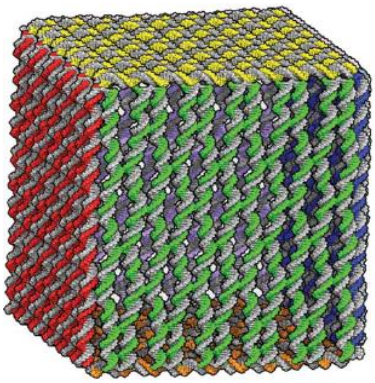


Figure 1.4. *Molecular model of six DNA sheets in a high order cubic structure.*[9]

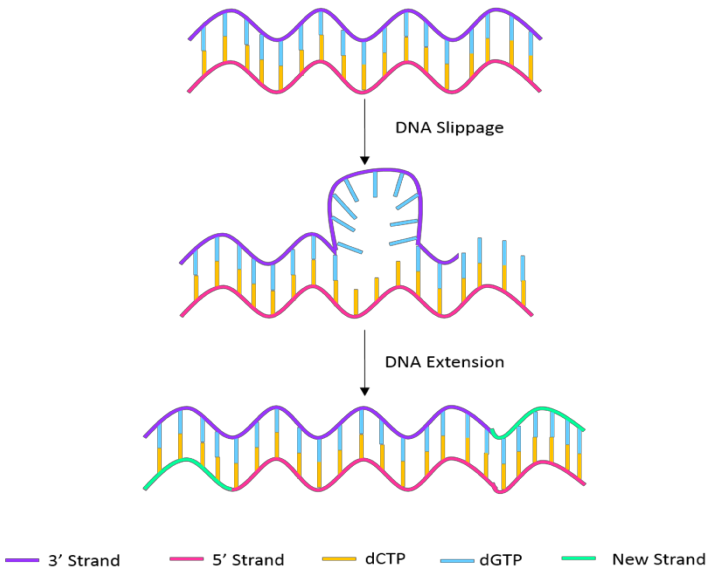
Much of the DNA nanomaterial synthesis relies on sources of DNA that are commercially available, be them long (Lambda) or short (oligos) in length. The possibilities of building new materials could be expanded further by the production of new kinds of DNA, where there is more control over the design of the sequence in DNA > 100 base pairs (bp). Therefore, this thesis investigates the use of DNA in two aspects; firstly, to synthesise new types of DNA and secondly, to use the newly designed DNA for surface fabrication of a diagnostic sensor. The following sections review the current state of DNA synthesis, common DNA modifications and a brief history of DNA chip technology.

1.1.1 DNA Synthesis

DNA can be synthesised in two ways, either by solid support automated synthesis or by enzymatic methods. Beaucage and Carruthers[33] developed phosphoramidites to ease the synthesis of DNA oligos. Automation of their coupling, washing and decoupling steps enabled the commercial availability of oligos up to 100 bp in length.[33,34] This method uses a controlled pore glass (CPG) solid support with an immobilised phosphoramidite as the initial nucleoside. Step by step addition of each base is performed in the 3' to 5' direction until the desired oligo length is reached. The oligo is cleaved from the CPG surface and purified by filtration followed by reverse phase high performance liquid chromatography (HPLC). However, the oligo length is limited to ~ 100 bp, due to the diminishing yield over many synthetic steps. Oligos less than the persistence length are more rigid, so are beneficial building blocks for nanomaterial assembly but their uses are limited on their own. The addition of internal modifications by automated synthesis to the oligo is possible, although the user's choice is commercially limited, unless they are modified in-house by the user.[35,36]

Enzymatic methods were, and still are, frequently used biologically to replicate DNA by primer/template polymerase chain reaction (PCR) for sensing purposes rather than using it as a nanomaterial. However, the generation of DNA as a material has become more popular with the exploitation of DNA enzymatic synthesis methods, such as slippage and rolling circle amplification (RCA).

The slippage reaction occurs when the DNA polymerase approaches the direct repeat during replication.[37–39] The polymerase delays replication by being briefly released from the template strand, allowing the new strand to separate from the template and match with another direct repeat. The DNA reorganises itself on the template strand and continues replication, **Scheme 1.2**.

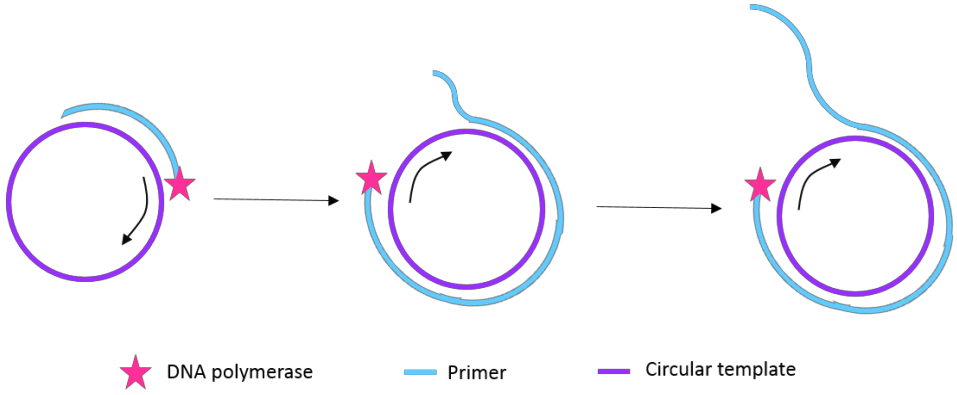


Scheme 1.2. *Process of the slippage reaction. Hairpin loop forms during the DNA slip and reanneals with extension of the strands.*[37]

The most commonly used polymerase for slippage is the the *E. coli* DNA polymerase I Klenow exonuclease minus (exo$^-$) fragment which is able to extend short oligos up to 10,000 bp.[39,40] Little preparation is required for this simple process; however, choosing a precise product length is time consuming and the sequence complexity is limited as the strand must be able to form hairpin loops in order to enable a slip.

RCA allows for the amplification of a short DNA primer by a circular template to generate long ssDNA by an isothermal, enzymatic process.[41] The DNA primer hybridises with a circular ssDNA template and the DNA polymerase copies the circular template by adding

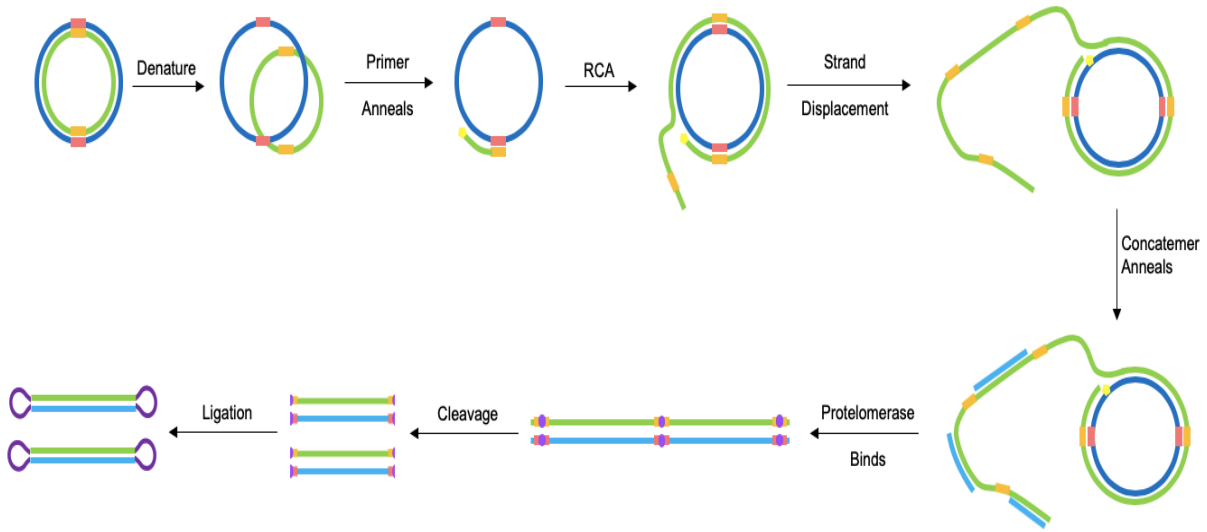
the relevant nucleotides. The ssDNA unwinds from the circular DNA template by strand displacement during polymerase replication, **Scheme 1.3**.[42] As a final step, a DNA restriction enzyme can be introduced to cut the long ssDNA into short fragments which is useful for the amplification of short DNA sequences for further applications.[43]



Scheme 1.3. *RCA amplification process.*[41,42]

There are advantages of using RCA as it is conducted at a constant temperature which is close to room temperature and does not require advanced instrumentation; however, the user is limited in length by the processivity of the DNA polymerase.[42] Furthermore, the resulting strands are single stranded, and there is no control over the design of the sequence as the user is confined to commercially available templates. The ability for researchers to produce specially designed circular templates would open up more opportunities for controlled DNA material synthesis.

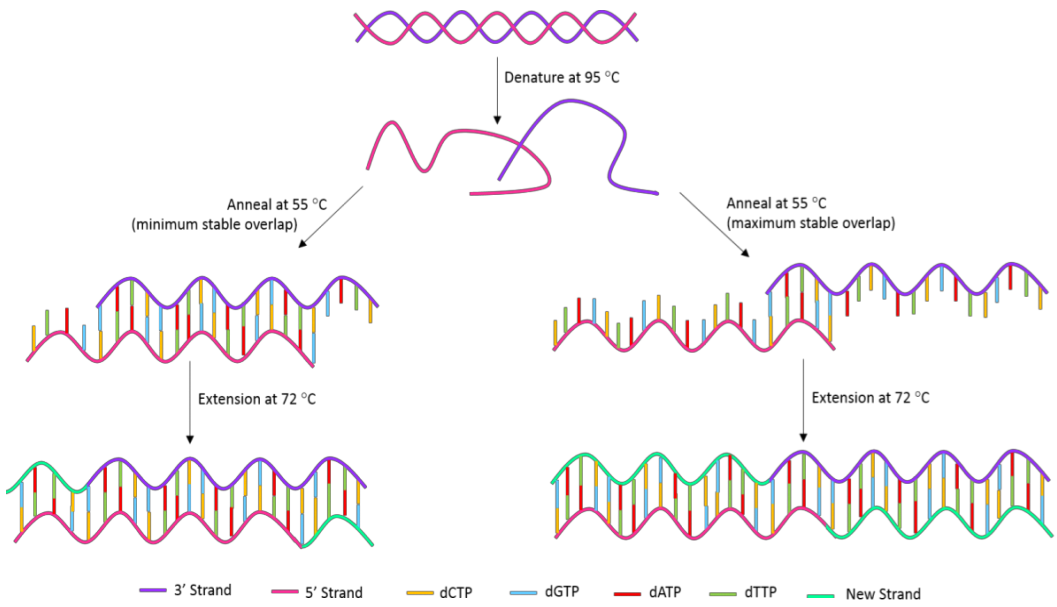
Touchlight Genetics Ltd. have advanced the RCA concept to produce covalently closed linear vector DNA, known as Doggybone DNA (dbDNA™) due to its structural design.[44–46] **Scheme 1.4** depicts the process of dbDNA™ production. Firstly, the circular DNA template is produced by ligation and denaturation. The DNA template contains protelomerase binding sites to enable cleavage of the DNA after RCA. The primer is bound to the template in the presence of a polymerase and RCA occurs. Strand displacement occurs and releases the extended DNA (extDNA) from the circular template and short DNA complement strands anneal to the extDNA to form long concatemers. Protelomerase binds to the amplified DNA and cleaves the concatemers into separate amplified strands. Ligation of the newly produced concatemer ends occurs to form dbDNA™.[46,47]



Scheme 1.4. *RCA to produce Doggybone DNA (dbDNA™).*[46]

This method has shown promise for the development of DNA vaccines which encode cancer antigens.[45,46] However, despite its ability to rapidly manufacture long and difficult sequences of DNA enzymatically, the DNA produced is single stranded unless concatemers are hybridised to the amplified DNA to attain a double strand.

More recently, Whitfield *et al.*[48,49] developed a new enzymatic method of DNA synthesis. The Whitfield method affords sequence control as it enables the user to design repeat DNA sequences which are specific to their needs and the synthesis allows for the introduction of base modifications if the user desires.[50] The enzymatic method draws inspiration from previous extension and amplification techniques to produce a PCR based heat-cool cycle extension method, **Scheme 1.5**.



Scheme 1.5. *Whitfield enzymatic extension process. dsDNA is denatured at 95 °C to form ssDNA which is annealed at 55 °C to allow a shift in the repeat sequence to occur. The exposed base pairs are filled in at 72 °C and process is repeated for the desired number of cycles.*[48,49]

The method consists of using heat-cool cycles to extend short repeat unit oligos of around 20 bases, known as oligoseeds. The oligoseed must consist of at least two repeating units to enable extension, while ensuring a minimum complementary stable overlap of 8 bp is maintained. Initially, the duplex is melted at 95 °C to form two single strands and the complementary regions anneal upon cooling to 55 °C. This produces "shifted" duplexes with 3' single stranded overhangs. The shifts can form a maximum or minimum overlap and their theoretical length can be calculated from Equation 1 and 2, respectively.

$$Minimum = original\ length + (y\ x\ n) \qquad \textbf{Equation 1}$$

$$Maximum = original\ length\ x\ 2^{y} - [(min.\ stable\ overlap\ x\ 2^{y}) - min.\ stable\ overlap] \qquad \textbf{Equation 2}$$

Where; n = number of base pairs in the repeat unit and y = number of cycles

After annealing, the sample is incubated at 72 °C for the polymerase catalysed extension of the 3' overhangs to occur. The heat-cool cycle described above is repeated until the desired length of DNA is obtained.

Z3 exo$^-$ which was developed by Jozwiakowski *et al.*[51] was the polymerase of choice for Whitfield *et al.*[48] during this research. Z3 is a *Thermococcus gorgonaruis* polymerase which belongs to the Family B polymerase group, where the primary function is to

accurately and efficiently replicate DNA.[52] Jozwiakowski *et al.*[51] found that Z3 has good thermal stability and is believed to be a more flexible polymerase, allowing it to incorporate more complex bases as well as read more complex sequences. In this Family of polymerases, the fingers domain connects with the deoxynucleotide triphosphates (dNTPs) and experiences a catalytic conformational change, requiring the correct base pairing between dNTP and template. Since Jozwiakowski *et al.*[51] reported that low-fidelity Pols may contain active sites which accommodate base pair mismatches, they believed that a higher tolerance level of unnatural base pairs would be accommodated by using low fidelity fingers (Pol ζ). Therefore, Z3 was designed to include Pol ζ fingers and a processivity enhancer (Sso7d) with the exonuclease carboxylates converted to alanine, in order to disable the 3' to 5' proofreading activity. Thermostable exo$^-$ DNA polymerases which are commercially available can also be used in the Whitfield enzymatic extension method.[49]

The Whitfield enzymatic extension approach affords a flexible method for producing long DNA sequences with multiple repeating units. A large variety of oligoseed sequences can be used with this method so there are few limitations to the user's sequence selection and complexity, unlike the other synthesis methods previously described. Although this method advances the DNA enzymatic extension methods, it still only affords some control over the length of DNA based on the heat-cool cycle lengths as a broad distribution of lengths is always produced and specific ranges of DNA size is not attainable, **Figure 1.5**.[48]

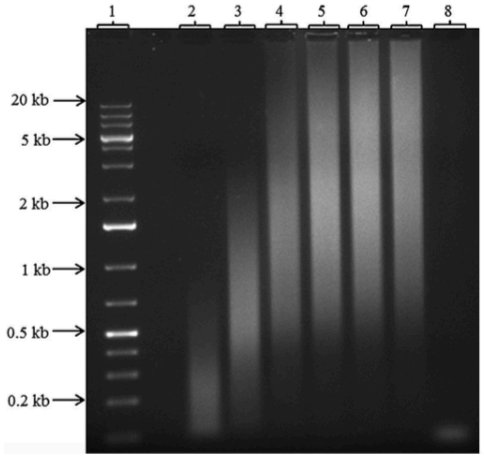
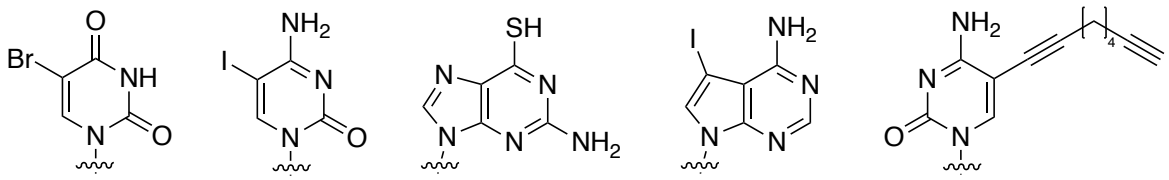


Figure 1.5. *0.3 % agarose gel of the extension products produced from the Whitfield enzymatic extension heat-cool cycling. Lane 1: DNA ladder; lane 2: 5 cycles; lane 3: 10 cycles; lane 4: 15 cycles; lane 5: 20 cycles; lane 6: 25 cycles; lane 7: 30 cycles; lane 8: no polymerase (control).*[48]

One other advantage of using DNA as a nanoscale building block is that it can be easily adapted by either base or backbone modification. Previous work has exploited DNA modifications for metal binding as the modifications can be incorporated by automated synthesis provided that the modified phosphoramidite is commercially available or can be synthesised.[53–55] Previously, research into on-column modification of an oligo has been reported by Lunn *et al.*,[55] where selective conversion of the G base to a thio-G base yielded modified oligos. The introduction of the thiol modification caused destabilisation of the duplex structure, however cadmium (II) coordination, directed by the thio-G, restabilised the duplex.

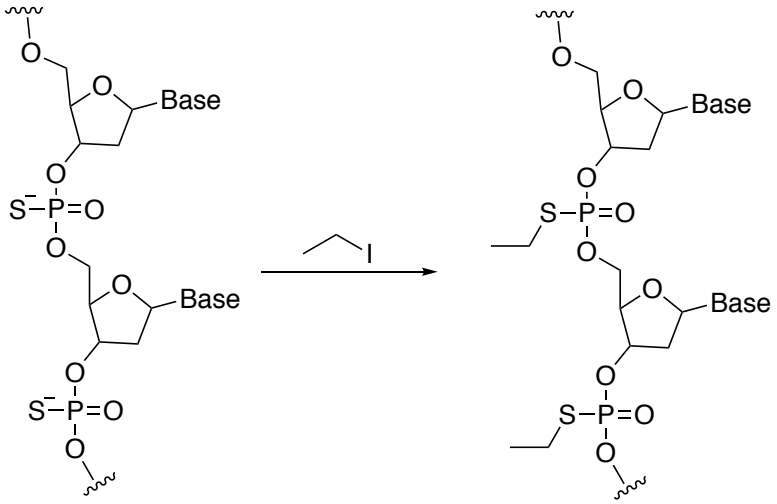
The incorporation of modified bases by automated synthesis is limited by the availability of modified phosphoramidites and the oligo length which can be achieved. Enzymatic fabrication of modified DNA is desirable for the study of DNA mutations, which occur naturally through oxidation or methylation of the base. These DNA mutations interfere with the cell replication and transcription processes and can lead to cancer and disease development. Kotlyar *et al.*[56] and Whitfield *et al.*[50] have reported that enzymatic methods can be used to produce modified DNA > 100 bp. Of note, Whitfield managed to synthesise base modified DNA of 500 bp which consisted of multiple modified bases such as; 5-bromo-deoxyuracil (5-Br-dU), 5-iodo-deoxycytidine (5-I-dC), 6-thiol-deoxyguanine (6-S-dG), 7-deaza-7-iodo-deoxyadenine (7-deaza-7-I-dA) and 5-octyne-deoxycytidine (5-C_8-alkyne-dC), **Scheme1.6**.[50]



Scheme 1.6. *Base modifications incorporated into the DNA by the Whitfield enzymatic extension method.*[50]

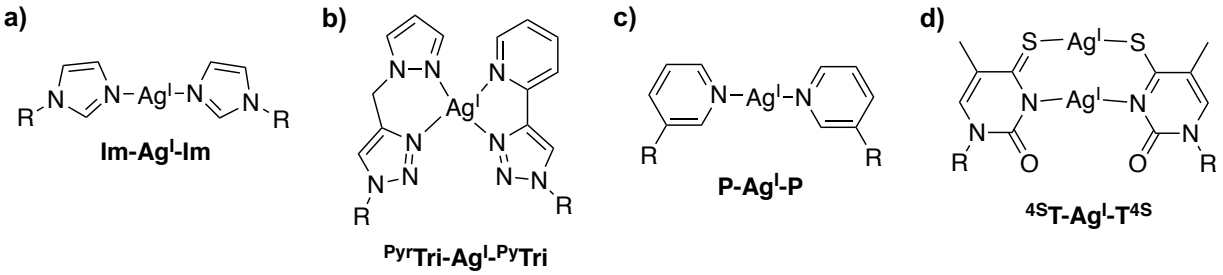
The enzymatic method used by Kotlyar *et al.*[56] produced backbone modified DNA where the phosphate groups were replaced with phosphorothioate groups, introducing thiol functionality to the backbone. Howorka's group furthered the modification of the phosphorothioate backbone by the addition of an alkyl group, **Scheme 1.7.**[57,58] This

makes the oligo hydrophobic by masking the negative charge on the backbone and is useful for applications in bio-sensing and controlled transmembrane transport.[57]



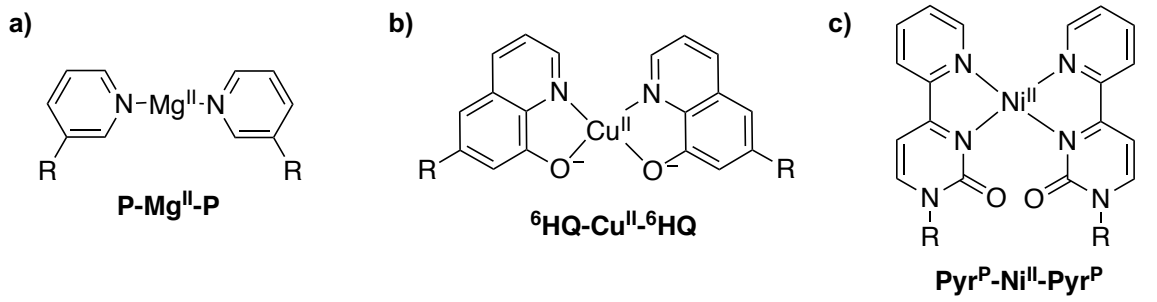
Scheme 1.7. *Modification of the phosphorothioate DNA backbone by alkylation.*[57,58]

Alternatively, the natural base can be replaced with an artificial base where the base pairing is metal-mediated[59] Shionoya's group first reported this system in 1999 where the nucleobase was modified with o-phenylenediamine to produce an artificial base that base-paired in the presence of palladium (II).[60] Over the span of this research, one of the most investigated artificial nucleobases with metal-mediated base pairing is the imidazole-silver(I)-imidazole system (Im-Ag^{I}-Im), **Scheme 1.8**.[61] This structure showed that the B-DNA geometry was maintained despite the incorporation of metal-mediated bases.[53] Alongside the imidazole system, Ag^{I} has been shown to enable metal-mediated base pairing between many other artificial bases including click nucleosides,[62] pyridine[63] and thiols[64] to name a few.



Scheme 1.8. *Artificial nucleobase mediated base pairing, where R = DNA.* ***a)*** *imidazole – Im-Ag^{I}-Im*[61] ***b)*** *click nucleosides – ^{Pyr}Tri- Ag^{I}-^{Pyr}Tri*[62] ***c)*** *pyridine – P-Ag^{I}-P*[63] ***d)*** *thiol - ^{4S}T- Ag$^{I}_{2}$-^{4S}T.*[64]

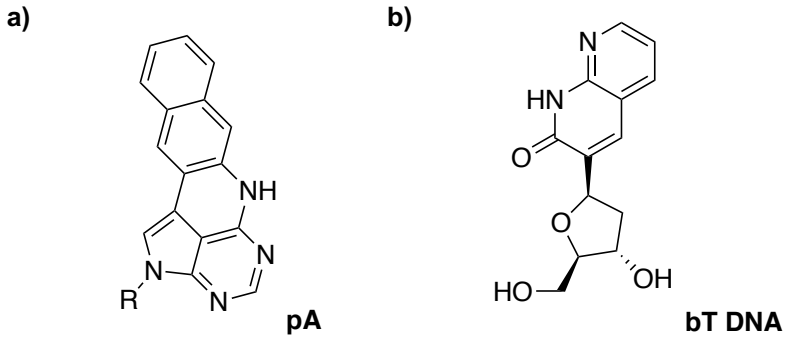
Mercury (II), copper (II) and nickel (II) are examples of other metal ions which can be accepted into the DNA duplex due to their ability to form planar metal complexes. All of these metal ions have been proven to induce metal-mediated base pairing as shown in **Scheme 1.9**.[65–67]



Scheme 1.9. *Metal-mediated base pairing induced by* ***a)*** *mercury (II)* ***b)*** *copper (II)* ***c)*** *nickel (II), where R = DNA.*[65–67]

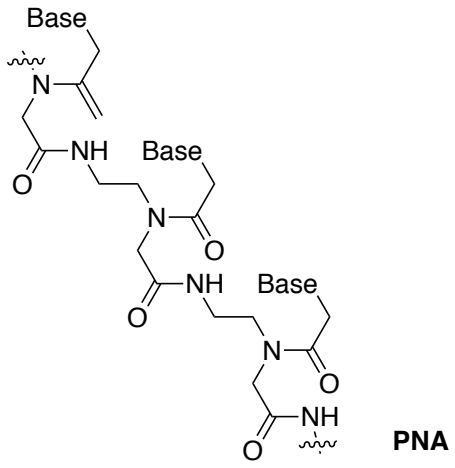
Moreover, Urata *et al.*[68] reported that these unnatural bases can be incorporated into the DNA sequence by using a polymerase. The artificial base was inserted opposite the natural T base in the presence of mercury (II) and the DNA continued to extend beyond the modification.

The natural bases can also be replaced by nucleosides which are modified to contain intrinsic fluorescence, known as fluorescent base analogues (FBA).[69,70] Incorporation of FBAs into the duplex structure can retain the hydrogen bonding of the natural bases and cause minimal distortion of the DNA structure.[71–73] The ability of these bases to position the fluorescence capabilities inside the DNA structure offers several advantages over external fluorophores which are covalently attached; less bulky, greater accuracy in position and orientation and the ability to label near the area of interest.[69] Pentacyclic adenine (pA)[69] and bicyclic thymidine (bT)[70] are good examples of FBAs developed by Tom Brown's group, **Scheme 1.10**. They can pair with their complementary bases well, which enables them to be incorporated enzymatically, as well as chemically on a solid support. FBAs have found applications in microenvironment sensing and proved useful in understanding the structure and dynamics of nucleic acids.[71–75]



Scheme 1.10. *Examples of fluorescent base analgoues (FBAs), where R = DNA.* ***a)*** *pentacyclic adenine (pA)* ***b)*** *bicyclic thymidine (bT).*[69,70]

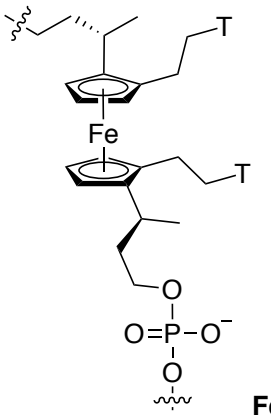
Instead of replacing the DNA bases, the phosphate backbone can be changed to introduce different moieties. One of the first examples of backbone replacement was the development of polyamide nucleic acid (PNA).[76] The artificial DNA uses aminoethylglycyl units to link the thymine bases, **Scheme 1.11**. The resulting oligo consisting of a polyamide backbone retains hybridisation with its complementary bases. The PNA hybridises to its complement strand in dsDNA via strand displacement due to the high stability of the PNA-DNA structures.



Scheme 1.11. *DNA backbone replacement with polyamide nucleic acid (PNA).*[76]

The Tucker group also showed that the phosphate backbone could be replaced to produce a ferrocene nucleic acid oligo (FcDNA).[77] The FcDNA incorporates the metal into the backbone by replacing the sugar-phosphate-sugar groups with ferrocene (cyclopentadienyl-iron-cyclopentadienyl), **Scheme 1.12**. Thymine functionalised ferrocene units are incorporated into the oligo by automated synthesis on a CPG support. Water solubility is ensured by inclusion of a negatively charged phosphodiester unit to

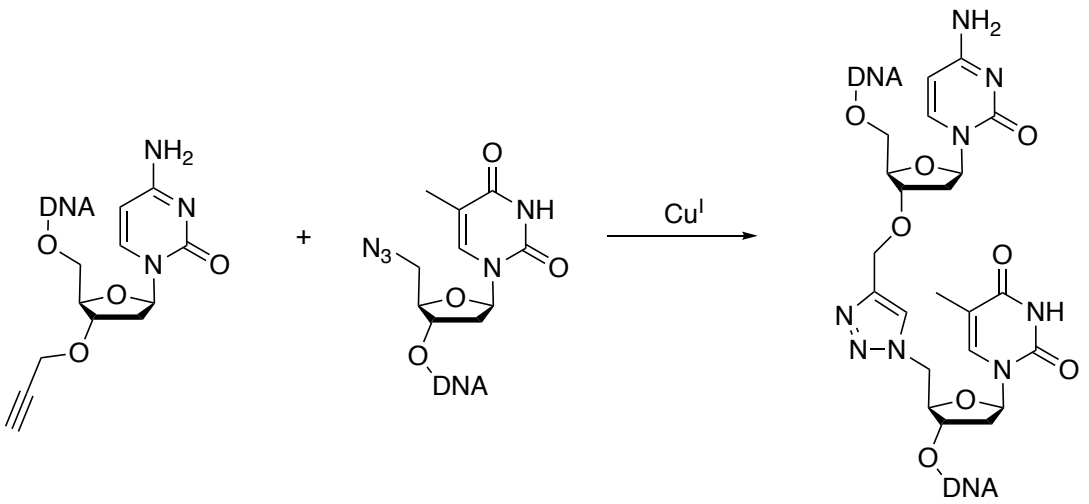
bridge the ferrocenes. Despite the replacement of the phosphate backbone with FcDNA, the DNA duplex structure is maintained.



FcDNA

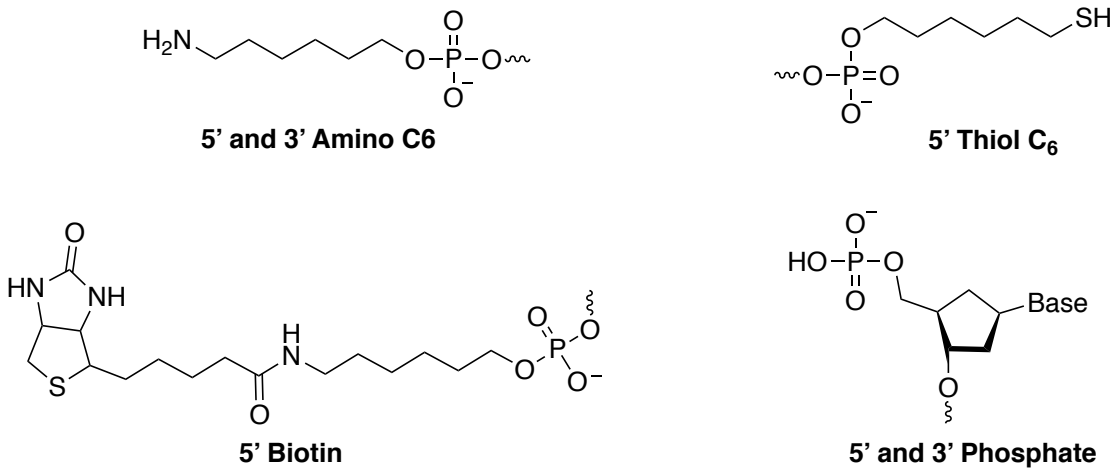
Scheme 1.12. *Replacement of the DNA backbone with ferrocene nucleic acid (FcDNA).*[77]

The generation of chemically ligated oligos by Tom Brown's group is another example of the replacement of the phosphate backbone.[78–80] The group utilised the CuAAC 'click' reaction, developed by Sharpless[81] and Meldal,[82] to form a thiazole backbone linkage.[78] El-Sagheer *et al.*[78] reported that triazole linkages could be generated by using the CuAAC reaction with a 3'-alkyne oligo, a 5'-azide oligo and a bifunctional 3'-alkyne-5'-azide oligo, **Scheme 1.13**. The 100 bp oligos could be ligated to form a 300 bp DNA strand which contained two thiazole linkages. These artificial DNA strands were also found to be compatible with PCR amplification as the sequencing of the PCR product confirmed that a true copy of the original template was produced.



Scheme 1.13. *CuAAC 'click' reaction of alkyne and azide oligos to produce DNA with a thiazole linked backbone.*[78,79]

Oligos can also accept modifications at the 5' and 3' termini. The 5' and 3' terminal ends generally consist of hydroxyl groups or phosphate groups depending on the cleavage from the solid support. These modifications tend to be achieved through C_3 and C_6 linkers to give the user a choice of terminal functional groups; thiol, biotin, amine, aldehyde *etc.*, **Scheme 1.14**. There are many terminal functionalities which are commercially available and have been exploited for the assembly of monolayers on a surface. DNA surfaces of this nature have been heavily exploited for the modification of gold nanoparticle (AuNP) surfaces for drug delivery,[83] conductive materials[84] and DNA chips for gene sensing.[85] The following section further details the generation of DNA immobilised on surfaces, with a focus on DNA biosensors.



Scheme 1.14. *Commercial 5' and 3' terminal oligo modifications.*

1.1.2 DNA Surfaces

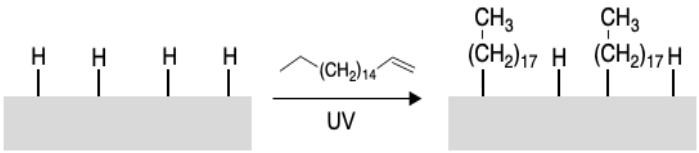
DNA nanotechnology provides a positive approach in the determination of biomolecular recognition events due to its programmable assembly. DNA surfaces with a high order organisation of the biomolecules can be fabricated by DNA self-assembly on the surface.[86,87] These surfaces are classed as a biomimetric analytical tool for applications in drug discovery,[88] bio-sensing[89] and diagnostics.[90] Despite the current surface limitations of sensitivity, specificity and reproducibility, DNA surfaces are expected to be an important addition in medical diagnosis and the determination of a treatment route for the patient, due to the increase in identifiable biomarkers for numerous diseases.[91]

The DNA surface fabrication mechanism can vary depending on the sensor application; however, they all require the immobilisation of a probe DNA strand on a solid support, to

which the target DNA strand can bind to. The DNA can be immobilised onto various solid supports which act as a surface. Silicon (Si), flat gold (Au), AuNPs and glass are all examples of surfaces which have been exploited, with each having their own preferred chemistry for DNA attachment.

Silicon surfaces are used for DNA field-effect transistors as they are atomically flat so allow for uniform DNA attachment and electronic characterisation of the surface.[92,93] The same chemistry can be utilised on porous silicon, a large surface area substrate, and enables spectroscopic characterisation of the surface, *e.g.* FTIR.[94] DNA monolayers on silicon have also been immobilised on single crystal and silicon nanoparticles.[95] In all of these types of silicon substrates, a chemical bifunctional linker is required to enable DNA immobilisation. A wide range of linkers have been employed, utilising either hydrosilylation or siloxane chemistry.

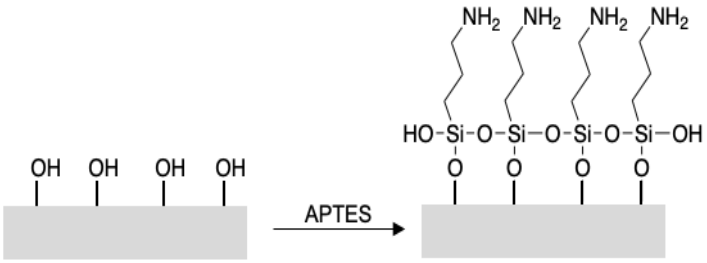
Hydrosilylation surface chemistry requires a hydrogen terminated substrate, **Scheme 1.15**, where no oxide is present. Unsaturated compounds, such as alkynes, alkenes, amines and aldehydes can react with the surface to form a monolayer bound through Si-C, Si-O or Si-N.[96–98] Therefore, these monolayers are stable to high temperatures (120 - 200 °C) and ultra-violet (UV) treatment.[99]



Scheme 1.15. *Hydrosilylation surface modification with 1-octadecene by UV treatment.*[97,99]

Siloxane surface modification has been successful on a wide variety of surfaces; glass, silicon, plastics, metals and metal oxides.[100] Siloxane chemistry on silicon exploits the native oxide layer, therefore removing the need for potentially harmful preparation steps. Typical silane linkers contain one to three hydrolysable groups, with the general structure $(MeO_3)Si(CH_2)_nX$,[101] which can react with the surface hydroxyl groups to give a (-O-Si-C-) attached siloxane monolayer, **Scheme 1.16**.[102] The tail groups of these linkers (represented by X in the general structure) provide the desired chemical moiety for DNA oligo attachment. The siloxane surfaces are thermally stable which is useful in DNA sensor fabrication where elevated temperatures, dehybridisation of the surface at 95 °C, are

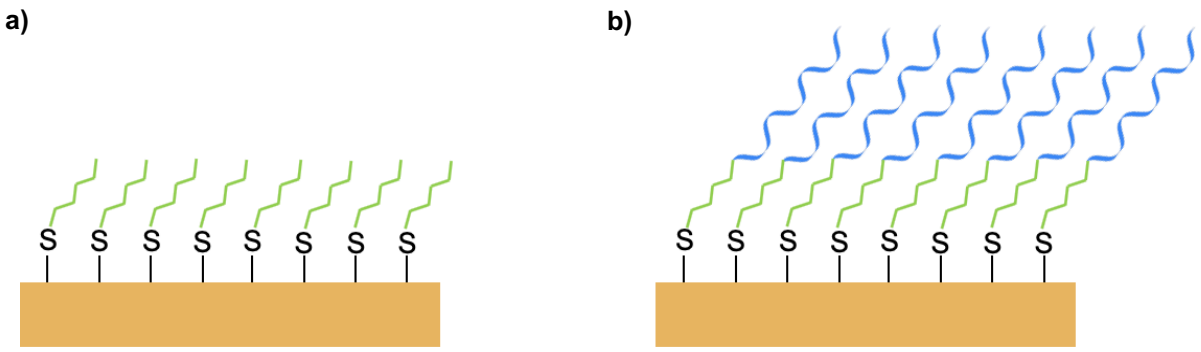
required.[103] Target DNA strands can then be rehybridised to the immobilised probe sequences.



Scheme 1.16. *Siloxane modification of a surface using (3-aminopropyl)triethoxysilane (APTES).*[101]

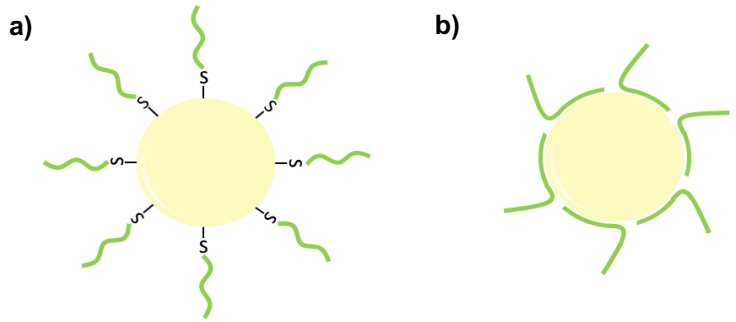
Gold surfaces have been utilised to produce DNA sensors as they can form well-ordered self-assembled monolayers and can be characterised electronically due to possessing conductive properties. By using gold, the DNA can be attached directly to the surface without linker deposition by taking advantage of forming strong covalent Au-S bonds. These bonds are formed by incorporating a thiol modification on the oligo at either the 5' or 3' strand terminus during automated synthesis.

Bain *et al.*,[104] reported that long chain alkanethiols, [SH-$(CH_2)_n$-X], in solution can adsorb onto gold surfaces to produce monolayers. These monolayers were densely packed and highly ordered as the thiol coordinated to the surface and the polymethylene chains packed in a tilted conformation to the surface, **Scheme 1.17**. This surface modification can be applied to the attachment of DNA as the X group can be in the form of a DNA oligo to create DNA immobilised surfaces. The alkane group is an advantageous addition as it increases the spacing of the DNA from the surface which eases the hybridisation of the target DNA by reducing steric hinderance resulting from the surface.



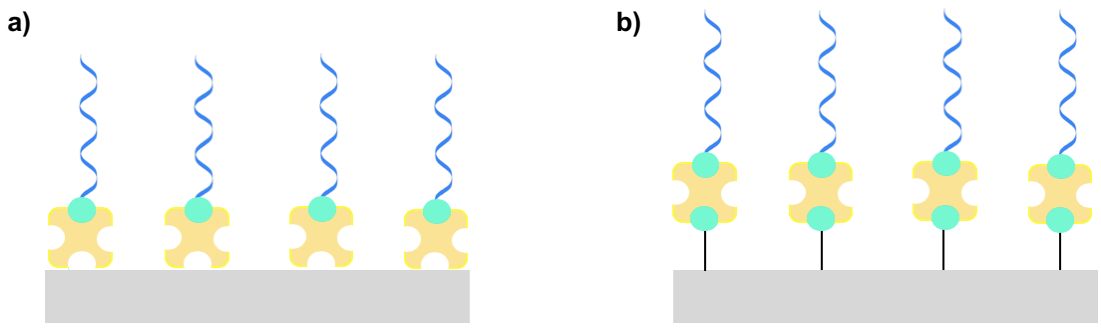
Scheme 1.17. *Alkanethiol gold surface modification* ***a)*** *without DNA*[104] ***b)*** *with DNA.*

AuNPs are advantageous in the fabrication of DNA sensors due to their high volume to surface ratio.[105] DNA-AuNP hybrid structures can be formed by terminal thiol functionalisation of the oligo[106–110] or by the interaction of the bases with the AuNP, **Scheme 1.18**.[109,111–113] These hybrids can be used for the production of bio-sensors[111] and nanoelectronics.[114]



Scheme 1.18. *DNA-AuNP hybrid fabrication.* ***a)*** *5' thiol functionalisation of the oligo[106–110]* ***b)*** *base interaction with the surface.[109,111–113]*

Streptavidin (SAv)-biotin surface modification is another method to produce DNA sensors. Biomolecular modification of the silica substrates is achieved by the deposition of SAv which can bind through the hydroxyl groups via hydrogen bonding.[115] SAv contains four binding sites which can accept biotin molecules, creating one of the strongest biomolecular interactions, and can therefore, accept terminally modified biotin oligos producing well-ordered DNA monolayers. SAv-biotin surfaces may also be generated by attaching a biotin linker to the surface and then depositing the SAv. The SAv binds to the covalently attached biotin on the surface and generates a more stable modified substrate compared to streptavidin addition alone. Again, biotin oligos can be attached to the one of the other free binding sites on the SAv to generate a DNA surface.[116,117]



Scheme 19. Streptavidin-biotin modification methods for DNA surface fabrication. ***a)*** *Streptavidin deposition and biotin-oligo binding*[115] ***b)*** *biotin linker modification, streptavidin deposition and biotin-oligo binding.*[116,117]

Currently, DNA sensors utilise the different methods of surface modification described, to immobilise relatively short DNA lengths (~ 100 bp) with terminal modifications due to automated synthesis limitations. Although there has been some research into the surface modification of DNA lengths around that of the persistence length by the Ijiro group.[116] Furthermore, Bracha *et al.*[118] reported that DNA up to 1,000 bp had been immobilised on the surface and the DNA packing density and salt concentration was varied to determine the effect of the surface conditions on the DNA length. It could be advantageous to have long DNA on the surface to increase the number of complementary binding sites for the target DNA hybridisation. More binding sites increases the confidence in the binding signal, which in most cases is a fluorescent read-out produced by the fluorophore tagged DNA.

1.2 Project Aims

The assembly of nanoscale DNA materials has relied on the use of short oligos to produce DNA tiles and Holliday junctions, or long commercially available DNA, M13mp18 genomic DNA, for the fabrication of DNA origami. The automated synthesis of long DNA, > 100 bp, is problematic and hence a method to fabricate DNA that is a useful material for nanoscale assembly is desired.

Recently, enzymatic extension methods[38,41,46] have been developed to fabricate long DNA which can be used as a material, however, these methods do not offer user control over the sequence. The production of designer DNA where the sequence can be specially controlled by the user, easily incorporate modifications and be of user defined lengths would be advantageous in the generation of nanomaterials and sensing devices.

Chapter 2 will discuss the use of the Whitfield enzymatic method to produce size specific designer DNA lengths. Approaches to amplify specific DNA lengths of a user designed sequence are discussed here, in order to obtain sufficient DNA to be useful as a nanomaterial. This development would open up an additional route for the production of more specialised DNA-based nanomaterials.

Further demonstration of the Whitfield enzymatic extension method will be discussed in Chapter 3. The fundamentals of this method are implemented in the synthesis of a new type of DNA, ABCD-DNA, composed of two different sequences. The versatility of this method is demonstrated by the inclusion of modified bases within one of the two sequences. Therefore, ABCD-DNA is able to control the assembly of nanomaterials through preferential spatial modification sites.

DNA based sensing devices are commonly fabricated by the immobilisation of DNA on a surface. Chapter 4 will investigate the extension of an immobilised DNA probe sequence by the Whitfield enzymatic method to create a fluorescently enhanced sensing platform.

1.3 References

1. Watson, S. M. D., Pike, A. R., Pate, J., Houlton, A. & Horrocks, B. R. DNA-templated nanowires: Morphology and electrical conductivity. *Nanoscale* **6**, 4027–4037 (2014).
2. Seeman, N. C. & Kallenbach, N. R. Design of immobile nucleic acid junctions. *Biophys. J.* **44**, 201–9 (1983).
3. Stoltenberg, R. M. & Woolley, A. T. DNA-Templated Nanowire Fabrication. *Biomed. Microdevices* **6**, 105–111 (2004).
4. Liu, W., Zhong, H., Wang, R. & Seeman, N. C. Crystalline two-dimensional DNA-origami arrays. *Angew. Chem. Int. Ed Engl.* **50**, 264–7 (2011).
5. Rothemund, P. W. K. Folding DNA to create nanoscale shapes and patterns. *Nature* **440**, 297–302 (2006).
6. Wei, B., Dai, M. & Yin, P. Complex shapes self-assembled from single-stranded DNA tiles. *Nature* **485**, 623–626 (2012).
7. Douglas, S. M. *et al.* Self-assembly of DNA into nanoscale three-dimensional shapes. *Nature* **459**, 414–418 (2009).
8. Goodman, R. P., Berry, R. M. & Turberfield, A. J. The single-step synthesis of a DNA tetrahedron. *Chem Commun* 1372–1373 (2004).
9. Andersen, E. S. *et al.* Self-assembly of a nanoscale DNA box with a controllable lid. *Nature* **459**, 73–76 (2009).
10. Dong, L. *et al.* Synthesis, Manipulation and Conductivity of Supramolecular Polymer Nanowires. *Chem. – Eur. J.* **13**, 822–828 (2007).
11. Al-Mahamad, L. L. G., El-Zubir, O., Smith, D. G., Horrocks, B. R. & Houlton, A. A coordination polymer for the site-specific integration of semiconducting sequences into DNA-based materials. *Nat. Commun.* **8**, 720 (2017).
12. Borovok, N., Molotsky, T., Ghabboun, J., Porath, D. & Kotlyar, A. Efficient procedure of preparation and properties of long uniform G4–DNA nanowires. *Anal. Biochem.* **374**, 71–78 (2008).
13. Drummond, T. G., Hill, M. G. & Barton, J. K. Electrochemical DNA sensors. *Nat. Biotechnol.* **21**, 1192–1199 (2003).
14. Seeman, N. C. Nanomaterials Based on DNA. *Annu. Rev. Biochem.* **79**, 65–87 (2010).
15. Zadegan, R. M. & Norton, M. L. Structural DNA nanotechnology: from design to applications. *Int. J. Mol. Sci.* **13**, 7149–7162 (2012).

16. Zahid, M., Kim, B., Hussain, R., Amin, R. & Park, S. H. DNA nanotechnology: a future perspective. *Nanoscale Res. Lett.* **8**, 119–119 (2013).
17. Ke, Y., Castro, C. & Choi, J. H. Structural DNA Nanotechnology: Artificial Nanostructures for Biomedical Research. *Annu. Rev. Biomed. Eng.* **20**, 375–401 (2018).
18. Hays, J. B., Magar, M. E. & Zimm, B. H. Persistence length of DNA. *Biopolymers* **8**, 531-536 (1969).
19. Murphy, C. *et al.* Long-range photoinduced electron transfer through a DNA helix. *Science* **262**, 1025–1029 (1993).
20. Braun, E., Eichen, Y., Sivan, U. & Ben-Yoseph, G. DNA-templated assembly and electrode attachment of a conducting silver wire. *Nature* **391**, 775–778 (1998).
21. Cohen, H., Nogues, C., Naaman, R. & Porath, D. Direct measurement of electrical transport through single DNA molecules of complex sequence. *Proc. Natl. Acad. Sci.* **102**, 11589–11593 (2005).
22. Livshits, G. I. *et al.* Long-range charge transport in single G-quadruplex DNA molecules. *Nat. Nanotechnol.* **9**, 1040 (2014).
23. Woiczikowski, P. B., Kubař, T., Gutiérrez, R., Cuniberti, G. & Elstner, M. Structural stability versus conformational sampling in biomolecular systems: Why is the charge transfer efficiency in G4-DNA better than in double-stranded DNA? *J. Chem. Phys.* **133**, (2010).
24. Richter, J., Mertig, M., Pompe, W., Mönch, I. & Schackert, H. K. Construction of highly conductive nanowires on a DNA template. *Appl. Phys. Lett.* **78**, 536–538 (2001).
25. Watson, S. M. D., Wright, N. G., Horrocks, B. R. & Houlton, A. Preparation, Characterization and Scanned Conductance Microscopy Studies of DNA-Templated One-Dimensional Copper Nanostructures. *Langmuir* **26**, 2068–2075 (2010).
26. Ford, W. E., Harnack, O., Yasuda, A. & Wessels, J. M. Platinated DNA as Precursors to Templated Chains of Metal Nanoparticles. *Adv. Mater.* **13**, 1793–1797 (2001).
27. Hassanien, R. *et al.* Preparation and Characterization of Conductive and Photoluminescent DNA-Templated Polyindole Nanowires. *ACS Nano* **4**, 2149–2159 (2010).
28. Seeman, N. C. Nucleic acid junctions and lattices. *J. Theor. Biol.* **99**, 237–247 (1982).
29. Mao, C., Sun, W. & Seeman, N. C. Designed Two-Dimensional DNA Holliday Junction Arrays Visualized by Atomic Force Microscopy. *J. Am. Chem. Soc.* **121**, 5437–5443 (1999).

30. Rothemund, P. W. K. & Andersen, E. S. The importance of being modular. *Nature* **485**, 584 (2012).
31. Kuzuya, A. & Komiyama, M. Design and construction of a box-shaped 3D-DNA origami. *Chem Commun* 4182–4184 (2009).
32. Shih, W. M. & Lin, C. Knitting complex weaves with DNA origami. *Curr. Opin. Struct. Biol.* **20**, 276–282 (2010).
33. Beaucage, S. L. & Caruthers, M. H. Deoxynucleoside phosphoramidites—A new class of key intermediates for deoxypolynucleotide synthesis. *Tetrahedron Lett.* **22**, 1859-1862 (1981).
34. Caruthers, M. H. *et al.* Chemical synthesis of deoxyoligonucleotides by the phosphoramidite method. *Recomb. DNA Part E* **154**, 287–313 (1987).
35. Pike, A. R. *et al.* Metallocene–DNA: Synthesis, Molecular and Electronic Structure and DNA Incorporation of C5-Ferrocenylthymidine Derivatives. *Chem. - Eur. J.* **8**, 2891 (2002).
36. Pike, A. R. *et al.* Ferrocenyl-Modified DNA: Synthesis, Characterization and Integration with Semiconductor Electrodes. *Chem. - Eur. J.* **11**, 344–353 (2005).
37. Wells, R. D., Dere, R., Hebert, M. L., Napierala, M. & Son, L. S. Advances in mechanisms of genetic instability related to hereditary neurological diseases. *Nucleic Acids Res.* **33**, 3785–3798 (2005).
38. Petruska, J., J Hartenstine, M. & F Goodman, M. Analysis of Strand Slippage in DNA Polymerase Expansions of CAG/CTG Triplet Repeats Associated with Neurodegenerative Disease. *J. Biol. Chem.* **273**, 5204–10 (1998).
39. Kotlyar, A. *et al.* Synthesis of novel poly(dG)-poly(dG)-poly(dC) triplex structure by Klenow exo À fragment of DNA polymerase I. *Nucleic Acids Res.* **33**, 6515–6521 (2005).
40. Kotlyar, A. B., Borovok, N., Molotsky, T., Fadeev, L. & Gozin, M. In vitro synthesis of uniform poly(dG)-poly(dC) by Klenow exo- fragment of polymerase I. *Nucleic Acids Res.* **33**, 525—535 (2005).
41. Ali, M. M. *et al.* Rolling circle amplification: a versatile tool for chemical biology, materials science and medicine. *Chem Soc Rev* **43**, 3324–3341 (2014).
42. Zhao, W., Ali, M. M., Brook, M. A. & Li, Y. Rolling Circle Amplification: Applications in Nanotechnology and Biodetection with Functional Nucleic Acids. *Angew. Chem. Int. Ed.* **47**, 6330–6337 (2008).
43. Banér, J., Mendel-Hartvig, M., Nilsson, M. & Landegren, U. Signal amplification of padlock probes by rolling circle replication. *Nucleic Acids Res.* **26**, 5073–5078 (1998).

44. Walters, A. A. *et al.* Comparative analysis of enzymatically produced novel linear DNA constructs with plasmids for use as DNA vaccines. *Gene Ther.* **21**, 645 (2014).
45. Allen, A. *et al.* Linear doggybone DNA vaccine induces similar immunological responses to conventional plasmid DNA independently of immune recognition by TLR9 in a pre-clinical model. *Cancer Immunol. Immunother.* **67**, 627–638 (2018).
46. Patel, A. *et al.* Novel synthetic plasmid and Doggybone™ DNA vaccines induce neutralizing antibodies and provide protection from lethal influenza challenge in mice AU - Scott, Veronica L. *Hum. Vaccines Immunother.* **11**, 1972–1982 (2015).
47. Heinrich, J. *et al.* Linear closed mini DNA generated by the prokaryotic cleaving-joining enzyme TelN is functional in mammalian cells. *J. Mol. Med.* **80**, 648–654 (2002).
48. Whitfield, C. J., Turley, A. T., Tuite, E. M., Connolly, B. A. & Pike, A. R. Enzymatic Method for the Synthesis of Long DNA Sequences with Multiple Repeat Units. *Angew. Chem. Int. Ed.* **54**, 8971–8974 (2015).
49. Little, R. C., Whitfield, C. J., Tuite, E. M. & Pike, A. R. The Synthesis of Designer DNA. in *DNA Nanotechnology: Methods and Protocols* (ed. Zuccheri, G.) 11–21 (Springer New York, 2018).
50. Whitfield, C. J. *et al.* Self-Priming Enzymatic Fabrication of Multiply Modified DNA. *Chem. - Eur. J.* **24**, 15267–15274 (2018).
51. Jozwiakowski, S. K., Keith, B. J., Gilroy, L., Doherty, A. J. & Connolly, B. A. An archaeal family-B DNA polymerase variant able to replicate past DNA damage: occurrence of replicative and translesion synthesis polymerases within the B family. *Nucleic Acids Res.* **42**, 9949–63 (2014).
52. Jannasch, H. W., Wirsen, C. O., Molyneaux, S. J. & Langworthy, T. A. Comparative Physiological Studies on Hyperthermophilic Archaea Isolated from Deep-Sea Hot Vents with Emphasis on Pyrococcus Strain GB-D. *Appl. Environ. Microbiol.* **58**, 3472 (1992).
53. Takezawa, Y. & Shionoya, M. Metal-Mediated DNA Base Pairing: Alternatives to Hydrogen-Bonded Watson–Crick Base Pairs. *Acc. Chem. Res.* **45**, 2066–2076 (2012).
54. Tanaka, K. *et al.* Efficient Incorporation of a Copper Hydroxypyridone Base Pair in DNA. *J. Am. Chem. Soc.* **124**, 12494–12498 (2002).
55. Lunn, S. M. L. *et al.* Duplex Healing of Selectively Thiolated Guanosine Mismatches through a Cd2+ Chemical Stimulus. *ChemBioChem* **19**, 1115–1118 (2018).
56. Ghabboun, J. *et al.* Specific and efficient adsorption of phosphorothioated DNA on Au-based surfaces and electrodes. *Appl. Phys. Lett.* **91**, (2007).

57. Burns, J. R., Stulz, E. & Howorka, S. Self-Assembled DNA Nanopores That Span Lipid Bilayers. *Nano Lett.* **13**, 2351–2356 (2013).
58. Gut, I. G. & Beck, S. A procedure for selective DNA alkylation and detection by mass spectrometry. *Nucleic Acids Res.* **23**, 1367–1373 (1995).
59. Takezawa, Y., Müller, J. & Shionoya, M. Artificial DNA Base Pairing Mediated by Diverse Metal Ions. *Chem. Lett.* **46**, (2016).
60. Tanaka, K. & Shionoya, M. Synthesis of a Novel Nucleoside for Alternative DNA Base Pairing through Metal Complexation. *J. Org. Chem.* **64**, 5002–5003 (1999).
61. Johannsen, S., Megger, N., Böhme, D., Sigel, R. K. O. & Müller, J. Solution structure of a DNA double helix with consecutive metal-mediated base pairs. *Nat. Chem.* **2**, 229 (2010).
62. Richters, T., Krug, O., Kösters, J., Hepp, A. & Müller, J. A Family of "Click" Nucleosides for Metal-Mediated Base Pairing: Unravelling the Principles of Highly Stabilizing Metal-Mediated Base Pairs. *Chem. – Eur. J.* **20**, 7811–7818 (2014).
63. Tanaka, K., Yamada, Y. & Shionoya, M. Formation of Silver(I)-Mediated DNA Duplex and Triplex through an Alternative Base Pair of Pyridine Nucleobases. *J. Am. Chem. Soc.* **124**, 8802–8803 (2002).
64. Okamoto, I., Ono, T., Sameshima, R. & Ono, A. Metal ion-binding properties of DNA duplexes containing thiopyrimidine base pairs. *Chem Commun* **48**, 4347–4349 (2012).
65. Zhang, L. & Meggers, E. An Extremely Stable and Orthogonal DNA Base Pair with a Simplified Three-Carbon Backbone. *J. Am. Chem. Soc.* **127**, 74–75 (2005).
66. Tanaka, K. *et al.* Programmable self-assembly of metal ions inside artificial DNA duplexes. *Nat. Nanotechnol.* **1**, 190 (2006).
67. Switzer, C. & Shin, D. A pyrimidine-like nickel(ii) DNA base pair. *Chem. Commun.* 1342-1344 (2005). doi:10.1039/B415426F
68. Urata, H., Yamaguchi, E., Funai, T., Matsumura, Y. & Wada, S. Incorporation of Thymine Nucleotides by DNA Polymerases through T–HgII–T Base Pairing. *Angew. Chem. Int. Ed.* **49**, 6516–6519 (2010).
69. Bood, M. *et al.* Pentacyclic adenine: a versatile and exceptionally bright fluorescent DNA base analogue. *Chem. Sci.* **9**, 3494–3502 (2018).
70. Lawson, C. P. *et al.* Synthesis, oligonucleotide incorporation and fluorescence properties in DNA of a bicyclic thymine analogue. *Sci. Rep.* **8**, 13970 (2018).
71. Xu, W., Chan, K. M. & Kool, E. T. Fluorescent nucleobases as tools for studying DNA and RNA. *Nat. Chem.* **9**, 1043 (2017).

72. Sinkeldam, R. W., Greco, N. J. & Tor, Y. Fluorescent Analogs of Biomolecular Building Blocks: Design, Properties, and Applications. *Chem. Rev.* **110**, 2579–2619 (2010).
73. Wilhelmsson, L. M. Fluorescent nucleic acid base analogues. *Q. Rev. Biophys.* **43**, 159-183 (2010).
74. Matarazzo, A. & Hudson, R. H. E. Fluorescent adenosine analogs: a comprehensive survey. *Tetrahedron* **71**, 1627–1657 (2015).
75. D. W. Dodd and R. H.E. Hudson. Intrinsically Fluorescent Base-Discriminating Nucleoside Analogs. *Mini-Rev. Org. Chem.* **6**, 378–391 (2009).
76. Nielsen, P., Egholm, M., Berg, R. & Buchardt, O. Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide. *Science* **254**, 1497-1500 (1991).
77. Nguyen, H. V. *et al.* A ferrocene nucleic acid oligomer as an organometallic structural mimic of DNA. *Chem Commun* **48**, 12165–12167 (2012).
78. El-Sagheer, A. H. & Brown, T. A triazole linkage that mimics the DNA phosphodiester group in living systems. *Q. Rev. Biophys.* **48**, 429–436 (2015).
79. Kukwikila, M., Gale, N., El-Sagheer, A. H., Brown, T. & Tavassoli, A. Assembly of a biocompatible triazole-linked gene by one-pot click-DNA ligation. *Nat. Chem.* **9**, 1089 (2017).
80. Kumar, A. *et al.* Assembling gold nanoparticles in solution using phosphorothioate DNA as structural interconnects. *Curr. Sci.* **84**, 71–74 (2003).
81. Kolb, H. C., Finn, M. G. & Sharpless, K. B. Click Chemistry: Diverse Chemical Function from a Few Good Reactions. *Angew. Chem. Int. Ed.* **40**, 2004–2021 (2001).
82. Tornøe, C. W., Christensen, C. & Meldal, M. Peptidotriazoles on Solid Phase: [1,2,3]-Triazoles by Regiospecific Copper(I)-Catalyzed 1,3-Dipolar Cycloadditions of Terminal Alkynes to Azides. *J. Org. Chem.* **67**, 3057–3064 (2002).
83. Ghosh, P., Han, G., De, M., Kim, C. K. & Rotello, V. M. Gold nanoparticles in delivery applications. *Adv. Drug Deliv. Rev.* **60**, 1307–1315 (2008).
84. Kim, H. J., Roh, Y., Kim, S. K. & Hong, B. Fabrication and characterization of DNA-templated conductive gold nanoparticle chains. *J. Appl. Phys.* **105**, 074302 (2009).
85. Hong, L. *et al.* Hybridization conditions of oligonucleotide-capped gold nanoparticles for SPR sensing of microRNA. *Biosens. Bioelectron.* **109**, 230–236 (2018).
86. Lin, M. *et al.* Programmable Engineering of a Biosensing Interface with Tetrahedral DNA Nanostructures for Ultrasensitive DNA Detection. *Angew. Chem. Int. Ed.* **54**, 2151–2155 (2015).

87. Lin, M. *et al.* Electrochemical detection of nucleic acids, proteins, small molecules and cells using a DNA-nanostructure-based universal biosensing platform. *Nat. Protoc.* **11**, 1244 (2016).
88. Yan, Z. & Wang, J. Specificity quantification of biomolecular recognition and its implication for drug discovery. *Sci. Rep.* **2**, 309 (2012).
89. Nimse, S. B., Song, K., Sonawane, M. D., Sayyed, D. R. & Kim, T. Immobilization Techniques for Microarray: Challenges and Applications. *Sensors* **14**, 22208–22229 (2014).
90. Pei, H., Zuo, X., Zhu, D., Huang, Q. & Fan, C. Functional DNA Nanostructures for Theranostic Applications. *Acc. Chem. Res.* **47**, 550–559 (2014).
91. Ye, D., Zuo, X. & Fan, C. DNA Nanotechnology-Enabled Interfacial Engineering for Biosensor Development. *Annu. Rev. Anal. Chem.* **11**, 171–195 (2018).
92. Wong, H.-S. P. Beyond the conventional transistor. *IBM J. Res. Dev.* **46**, 13–168 (2002).
93. Pouthas, F., Gentil, C., Côte, D. & Bockelmann, U. DNA detection on transistor arrays following mutation-specific enzymatic amplification. *Appl. Phys. Lett.* **84**, 1594–1596 (2004).
94. Bateman, J. E., Eagling, R. D., Worrall, D. R., Horrocks, B. R. & Houlton, A. Alkylation of Porous Silicon by Direct Reaction with Alkenes and Alkynes. *Angew. Chem. Int. Ed.* **37**, 2683–2685 (1998).
95. Lie, L. H. *et al.* Immobilisation and synthesis of DNA on Si(111), nanocrystalline porous silicon and silicon nanoparticles. *Faraday Discuss.* **125**, 235–249 (2004).
96. Boukherroub, R., Morin, S., Bensebaa, F. & Wayner, D. D. M. New Synthetic Routes to Alkyl Monolayers on the Si(111) Surface1. *Langmuir* **15**, 3831–3835 (1999).
97. Effenberger, F., Götz, G., Bidlingmaier, B. & Wezstein, M. Photoactivated Preparation and Patterning of Self-Assembled Monolayers with 1-Alkenes and Aldehydes on Silicon Hydride Surfaces. *Angew. Chem. Int. Ed.* **37**, 2462–2464 (1998).
98. Stewart, M. P. & Buriak, J. M. Photopatterned Hydrosilylation on Porous Silicon. *Angew. Chem. Int. Ed.* **37**, 3257–3260 (1998).
99. Cicero, R. L., Linford, M. R. & Chidsey, C. E. D. Photoreactivity of Unsaturated Compounds with Hydrogen-Terminated Silicon(111). *Langmuir* **16**, 5688–5695 (2000).
100. Chrisey, L. A., Lee, G. U. & O'Ferrall, C. E. Covalent attachment of synthetic DNA to self-assembled monolayer films. *Nucleic Acids Res.* **24**, 3031–3039 (1996).

101. Henke, L., Piunno, P. A. E., McClure, A. C. & Krull, U. J. Covalent immobilization of single-stranded DNA onto optical fibers using various linkers. *Anal. Chim. Acta* **344**, 201–213 (1997).

102. Lowe, R. D., Pellow, M. A., Stack, T. D. P. & Chidsey, C. E. D. Deposition of Dense Siloxane Monolayers from Water and Trimethoxyorganosilane Vapor. *Langmuir* **27**, 9928–9935 (2011).

103. Chandekar, A., Sengupta, S. K. & Whitten, J. E. Thermal stability of thiol and silane monolayers: A comparative study. *Appl. Surf. Sci.* **256**, 2742–2749 (2010).

104. Bain, C. D. & Whitesides, G. M. Molecular-Level Control over Surface Order in Self-Assembled Monolayer Films of Thiols on Gold. *Science* **240**, 62–63 (1988).

105. Saha, K., Agasti, S. S., Kim, C., Li, X. & Rotello, V. M. Gold Nanoparticles in Chemical and Biological Sensing. *Chem. Rev.* **112**, 2739–2779 (2012).

106. Zikich, D., Borovok, N., Molotsky, T. & Kotlyar, A. Synthesis and AFM Characterization of Poly(dG)-poly(dC)-gold Nanoparticle Conjugates. *Bioconjug. Chem.* **21**, 544–547 (2010).

107. Nicewarner Peña, S. R., Raina, S., Goodrich, G. P., Fedoroff, N. V. & Keating, C. D. Hybridization and Enzymatic Extension of Au Nanoparticle-Bound Oligonucleotides. *J. Am. Chem. Soc.* **124**, 7314–7323 (2002).

108. Eidelshtein, G., Fattal, M., Avishai, G., Giannini, C. & Kotlyar, A. Preparation, Characterization and Manipulation of Conjugates between Gold Nanoparticles and DNA. *Nanomaterials* **6**, 167–175 (2016).

109. Carnerero, J. M., Jimenez-Ruiz, A., Castillo, P. M. & Prado-Gotor, R. Covalent and Non-Covalent DNA-Gold-Nanoparticle Interactions: New Avenues of Research. *ChemPhysChem* **18**, 17–33 (2017).

110. Sandström, P., Boncheva, M. & Åkerman, B. Nonspecific and Thiol-Specific Binding of DNA to Gold Nanoparticles. *Langmuir* **19**, 7537–7543 (2003).

111. Pei, H. *et al.* Designed Diblock Oligonucleotide for the Synthesis of Spatially Isolated and Highly Hybridizable Functionalization of DNA–Gold Nanoparticle Nanoconjugates. *J. Am. Chem. Soc.* **134**, 11876–11879 (2012).

112. Zhou, W., Wang, F., Ding, J. & Liu, J. Tandem Phosphorothioate Modifications for DNA Adsorption Strength and Polarity Control on Gold Nanoparticles. *ACS Appl. Mater. Interfaces* **6**, 14795–14800 (2014).

113. Zhang, X. & Hu, H. DNA molecules site-specific immobilization and their applications. *Cent. Eur. J. Chem.* **12**, 977–993 (2014).

114. Park, J.-Y. & Park, S.-M. DNA Hybridization Sensors Based on Electrochemical Impedance Spectroscopy as a Detection Tool. *Sensors* **9**, 9513–9532 (2009).

115. Korwin-Edson, M. L., Clare, A. G., Hall, M. M. & Goldstein, A. Biospecificity of glass surfaces: streptavidin attachment to silica. *J. Non-Cryst. Solids* **349**, 260–266 (2004).

116. Mitomo, H. *et al.* Preparation and Characterization of Double-Stranded DNA Brushes via Surface-Initiated Enzymatic Polymerization. *J. Nanosci. Nanotechnol.* **17**, 8995–9001 (2017).

117. Nakamura, S. *et al.* DNA Brush-Directed Vertical Alignment of Extensive Gold Nanorod Arrays with Controlled Density. *ACS Omega* **2**, 46 (2017).

118. Bracha, D., Karzbrun, E., Shemer, G., Pincus, P. A. & Bar-Ziv, R. H. Entropy-driven collective interactions in DNA brushes on a biochip. *Proc. Natl. Acad. Sci.* **110**, 4534-4538 (2013).

Chapter 2

Amplification of Enzymatically Extended Size Specific Designer DNA

This chapter is based on the following book chapter:

Little, R.C., Whitfield, C.J., Tuite, E.M. & Pike, A.R., The Synthesis of Designer DNA, *DNA Nanotechnology*, 2018, 11-21

Contents

2.1 Introduction

Recent enzymatic methods for the synthesis of DNA over 100 bp are attractive for nanomaterial applications and biosensing. However, as discussed in **Chapter 1**, a common drawback of these enzymatic approaches is the inability to produce large quantities of precise DNA lengths. This chapter focusses on improving one method in particular to meet this goal, the Whitfield enzymatic extension method.[1]

In the Whitfield method, short repeat unit oligoseeds are extended to produce DNA products of a large base pair length distribution from 20 bp - 20,000 bp as observed in **Figure 2.1a**. The aim of this research is to isolate and amplify individual DNA fragments of a specific base pair length, *e.g.* 100-200 bp, **Figure 2.1b**. The Whitfield method is appealing as it overcomes the handicap of the diminishing yields of automated DNA synthesis, while maintaining control over the sequence. The ability to generate size specific DNA > 100 bp cost-effectively, is of increasing demand within the field of DNA based nanomaterials.

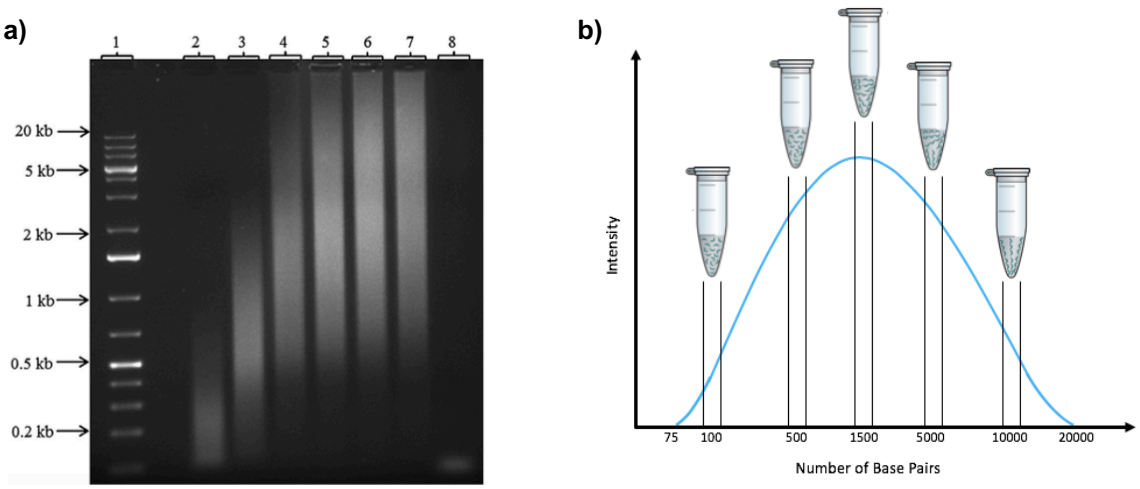


Figure 2.1. a) *0.3 % agarose gel of the extension products produced from the Whitfield enzymatic extension heat-cool cycling. Lane 1: DNA ladder; lane 2: 5 cycles; lane 3: 10 cycles; lane 4: 15 cycles; lane 5: 20 cycles; lane 6: 25 cycles; lane 7: 30 cycles; lane 8: no polymerase (control).*[1] **b)** *Production of size specific DNA by extraction from a wide distribution of extended DNA lengths. The removed base pair length of DNA can be user-defined within the limits of the DNA ladder use. For example, 100 bp, 500 bp, 1,500 bp, 5,000 bp and 10,000.*

The most efficient method for DNA size separation is by agarose or polyacrylamide gel electrophoresis. Many different techniques for the extraction of DNA from agarose gels have been published and include electroelution,[2] mechanical disruption and gel diffusion[3]

and dissolution of the gel by chaotropic agents.[4,5] It has been reported that it is difficult to remove high yields of DNA fragments and the recovered purity is low, which affects further enzymatic manipulation.[6]

In 1978, Tabak *et al.*[7] removed DNA from the agarose gel by the insertion of a trough filled with hydroxyapatite in front of the migrating DNA. The hydroxyapatite binds DNA which can then be removed by pipette after electrophoresis. The DNA was released from the hydroxyapatite by column chromatography using a Sephadex G50 column and phosphate buffer. This method recovered 80 % of the initial DNA, with losses attributed to incomplete removal of the solution from the trough and the DNA not fully eluting from the column.

A few years later, Dretzen *et al.*[8] proposed that DNA could be recovered during agarose gel electrophoresis by the migration of the DNA fragments onto strips of DEAE-cellulose paper inserted between the DNA bands visualised by ethidium bromide fluorescent staining. The DNA was recovered from the paper by shredding upon vortexing and incubation at 37 °C for 2 hours. The incubated solution was further manipulated by centrifugation and the DNA product precipitated with ethanol. This method worked well for fragments < 2,000 bp as 80 % recovery was achieved, however recovery was reduced to 60 % for longer fragments.

Gel excision has been routinely exploited due to the commercialisation of DNA extraction kits.[9–11] During the excision process the desired DNA fragments are identified based on a DNA ladder, before removal from the agarose using a cover slip and razor blade. The gel slice is placed into an Eppendorf and a commercially available gel extraction purification kit is used to purify the recovered DNA. Although gel extraction is an effective process,[12] a less intricate and time-consuming method is desirable.

Size exclusion HPLC has also been investigated as an alternative to the use of time-consuming agarose gels for the separation of different sized DNA fragments. Schmitter *et al.*[13] used size exclusion HPLC for the fast separation of DNA lengths between 75 – 10,000 bp. After elution and collection from the column, the DNA was precipitated by ethanol to yield biologically active DNA, unlike that produced by previous methods.

Kotlyar *et al.*[14,15] used size exclusion HPLC to recover different lengths of single stranded DNA. This was first described by the separation of poly(dG)-poly(dC) strands that had

been enzymatically synthesised to the same length by Klenow exonuclease minus (exo^-) fragment. The G strand was dehybridised from the C strand in 0.1 M sodium hydroxide and both strands were independently isolated by size exclusion HPLC. The G strands eluted earlier, however due to the similar retention time to the C strands, only the initial elution of G strands was pure. Therefore, only a < 5 % yield of separated G strands was obtained.

Using a similar approach, Kotlyar's group advanced this method to increase the final separated G strand yield to ~ 100 %,[16] by synthesising their long poly(dG) strands from a $poly(dG)_{10}$-$poly(dC)_{10}$ primer template, short $poly(dC)_{20}$ fragments and dGTP only. This produced one long poly(dG) strand and many short poly(dC) complement strands which were easier to separate by size exclusion HPLC. The difference in length allowed for longer retention times during elution and led to the increase in yield of the separated long poly(dG) strands.

All of the methods described above recovered well-defined DNA fragments either from gel electrophoretic methods or size exclusion HPLC. However, these methods have limitations in that they are time consuming and some require expensive equipment. Furthermore, in these cases the DNA was of a pre-defined length or the products were single stranded rather than double stranded. Therefore, it would be advantageous to have a rapid and versatile DNA size recovery method which produced dsDNA fragments at user defined lengths. This chapter describes the approach developed to obtain designer DNA of specific base pair lengths by agarose gels from DNA products synthesised by the Whitfield enzymatic extension method.[1,17]

2.2 Results and Discussion

Wide distributions of DNA lengths generated by the Whitfield enzymatic extension method were used to produce aliquots of size specific base pair length DNA by gel extraction. The concentrations of the size recovered aliquots could be less than 10 ng/μL and were required to be amplified to enable use in further DNA nanotechnology experimentation. Ligation techniques were employed to allow for copies of the target DNA to be made by qPCR. Here, varying DNA sequences (**Table 2.1**), ligation techniques and qPCR conditions were assessed for the production of over 500 μL of accurate target DNA at a concentration of 100 μM.

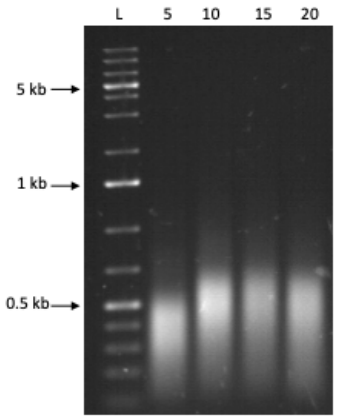
Table 2.1. *Oligoseed base sequences extended during the Whitfield enzymatic extension.*

Name	Sequence
$[AT]_{10}/[TA]_{10}$	5'-ATATATATATATATATATAT-3'/ 3'-TATATATATATATATATATA-5'
$[GC]_{10}/[CG]_{10}$	5'-GCGCGCGCGCGCGCGCGCGC-3'/ 3'-CGCGCGCGCGCGCGCGCGCG-5'

2.2.1 $[AT]_{10}/[TA]_{10}$ Oligo

Various heat-cool cycle lengths of the Whitfield enzymatic extension were investigated to determine the optimum growth length the oligoseed could achieve. The oligoseed consisted of dinucleotide repeats of dA and dT bases and were 20 bases in length, $[AT]_{10}/[TA]_{10}$ (**Table 2.1)**. The oligoseed was extended using *Thermococcus gorgonarius* polymerase B (Pfu-Pol) Z3 exo^{-18} for 5, 10, 15 and 20 heat-cool cycles performed in the PCR, where one heat-cool cycle was comprised of 95 °C for 30 seconds, 55 °C for 30 seconds and 72 °C for 120 seconds. The resulting product was purified using a commercial PCR purification kit and characterised by agarose gel and ultraviolet-visible (UV-Vis) spectroscopy, **Figure 2.2**.

a)



b)

No. of Cycles	Abs @ 260 nm	Concentration (ng/µL)
5	1.26	62.9
10	1.45	72.7
15	1.67	83.3
20	1.43	71.4

Figure 2.2. a) *Enzymatic extension with Tgo-Pol Z3 exo- and $[AT]_{10}/[TA]_{10}$ oligo seed for varying heat-cool cycle length. Agarose gel electrophoresis, lanes 1-4: extension products after 5, 10, 10 and 20 cycles, respectively and L = DNA ladder.* ***b)*** *UV-Vis absorbance of enzymatic extension products and concentration taken at 260 nm.*

A gradual increase in base pair length was evident on the agarose gel with respect to increasing cycle number, **Figure 2.2a**, however this difference was not very significant, particularly between 20 and 30 heat-cool cycles. The 260 nm peak in the UV-Vis trace was used to calculate the concentration of the DNA in the sample (one absorbance unit is equal to 50 ng/µL). The 260 nm absorbance data, **Figure 2.2b**, showed that the concentrations followed the expected trend, *i.e.* increasing absorbance at 260 nm value with increasing cycle numbers, until 15 cycles before the concentration decreased to 71.4 ng/µL. Overall, there was not a large difference observed in the concentration values.

Furthermore, the extended DNA (extDNA) lengths after 10 heat-cool cycles were shorter than expected in comparison to the calculated theoretical base pair lengths, **Table 2.2**, from the equations detailed in **Chapter 1**.

Table 2.2 *Theoretical maximum and minimum base pair lengths achieved during the Whitfield enzymatic extension.*

No. of Cycles	Theoretical Maximum (bp)	Theoretical Minimum (bp)
5	392	30
10	7080	40
15	17080	50
20	27080	60

Image J analysis software[19] was used to create a graphical representation of the gel image which allowed for more accurate determination of the mean length of the extDNA produced and the base pair length range of the distribution for each cycle set. Image J graph, **Figure 2.3a**, was obtained by using the intensity of the DNA ladder to mark the base pair lengths for use as the x-axis and matching the ladder intensity to the sample intensity by measuring the full DNA distribution in the agarose gel, **Figure 2.3b**.

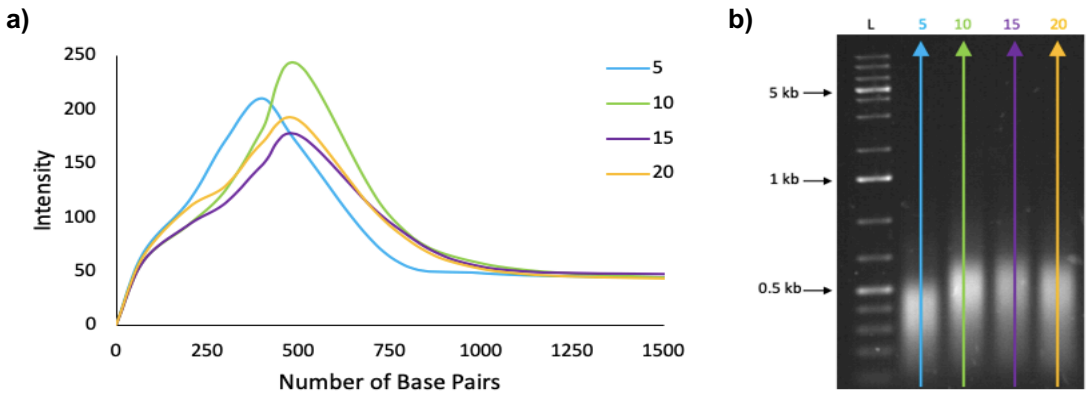
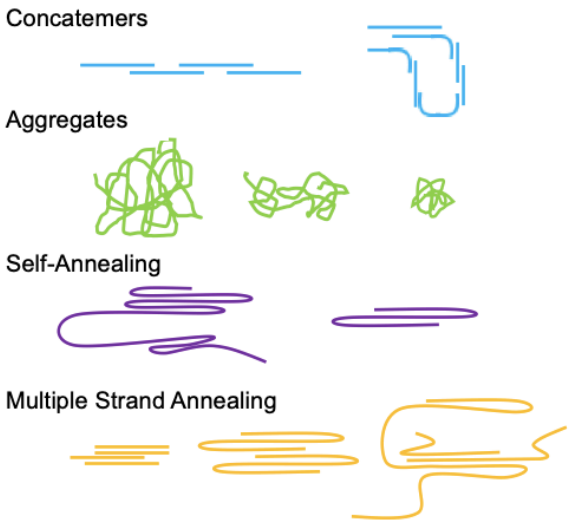


Figure 2.3. a) *Image J analysis of $[AT]_{10}/[TA]_{10}$ agarose gel after 5, 10, 15 and 20 enzymatic extension heat-cool cycles.* ***b)*** *agarose gel depicting the direction of intensity measurement which is transformed into the Image J plot profile graph.*

The peak maximum of each trace represented the average base pair length produced during the relevant number of heat-cool cycles. Since the increasing number of heat-cool cycles should generally increase the extended length of the DNA, it would be expected that the Image J peak maximum, which relates to the average number of base pairs, would increase along the x-axis (number of base pairs). It was clear that the peak maximums of the extDNA did not shift to longer base pair lengths after 5 cycles and that increasing the number of cycles was not generating DNA > 500 bp since the peak maximum for 10, 15 and 20 cycles did not increase any further along the x-axis of the graph. Furthermore, the base pair length ranges for the cycle lengths carried out in this experiment mildly exceeded 1,000 bp, which for 20 cycles was considered as very short compared to the calculated theoretical length, 27,080 bp.

The Whitfield extension synthesis was repeated many times for the $[AT]_{10}/[TA]_{10}$ oligoseed and it was apparent that some of the DNA samples had reached lengths of approximately 2,500 bp. On the basis of the theoretical lengths expected and the actual lengths seen with other oligoseeds in previous work,[20] the $[AT]_{10}/[TA]_{10}$ oligoseed produced relatively short DNA (~ 500 bp).

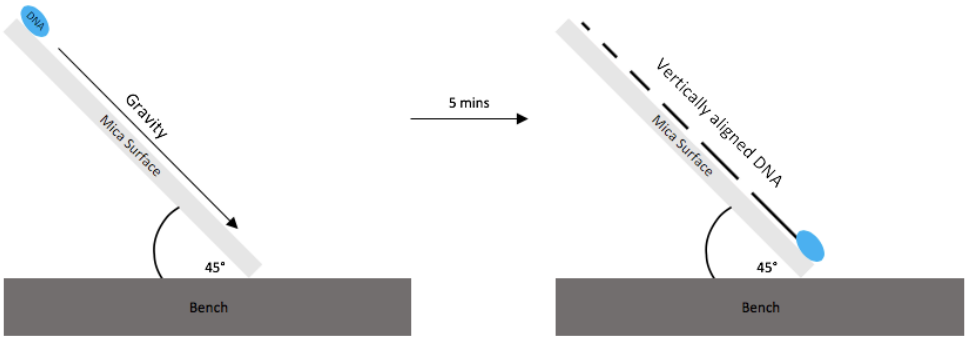
As previously discussed, the repetitive nature of the sequence may cause the extDNA to aggregate (anneal multiple strands to form a larger structure) or create concatemers (longer strands formed by shorter DNA strands annealing with each other, creating gaps in the sequence instead of fully formed dsDNA), through self-annealing or multiple strands annealing, **Scheme 2.1**.



Scheme 2.1. *DNA structures other than the linear form; concatemers, aggregates, self-annealing or multiple strand annealing.*

Formation of these structures would have affected the DNA's ability to extend creating DNA lengths shorter than expected. To determine the structure of the DNA and enable identification of aggregates or self-annealing, atomic force microscopy (AFM) was employed.

AFM is a technique that enables the user to image small samples with the resolution in the nanometre range[21,22] which makes it an ideal method for DNA analysis. The DNA of interest was fixed to a mica surface using divalent metal ions *e.g.* Mg^{2+}. The ions act as a bridge between the negatively charged DNA and the negatively charged surface, fixing the DNA onto the surface.[23]



Scheme 2.2. *Molecular combing of DNA across mica surface. Mica was fixed at 45° to the bench top and the DNA-ion solution was added along the top of the mica. Gravity pulled the droplet down the mica, and the DNA fixed in a gradient across the mica, from short to long in length.*[24]

Deposition of the DNA on the surface was performed via molecular combing,[24] **Scheme 2.2**. The mica surface was fixed at a 45° angle while the DNA-ion solution was added along the top of the mica surface. The DNA-ion solution flowed down the mica surfaces by gravity which allowed the DNA to stretch and align vertically where possible. DNA was fixed in a gradient pattern across the surface with the short DNA lengths at the top of the mica and long DNA lengths found at the bottom of the mica. AFM images of extended $[AT]_n/[TA]_n$ DNA were taken of a 4 ng/μL sample on freshly cleaved mica after 20 heat-cool cycles, **Figure 2.4**.

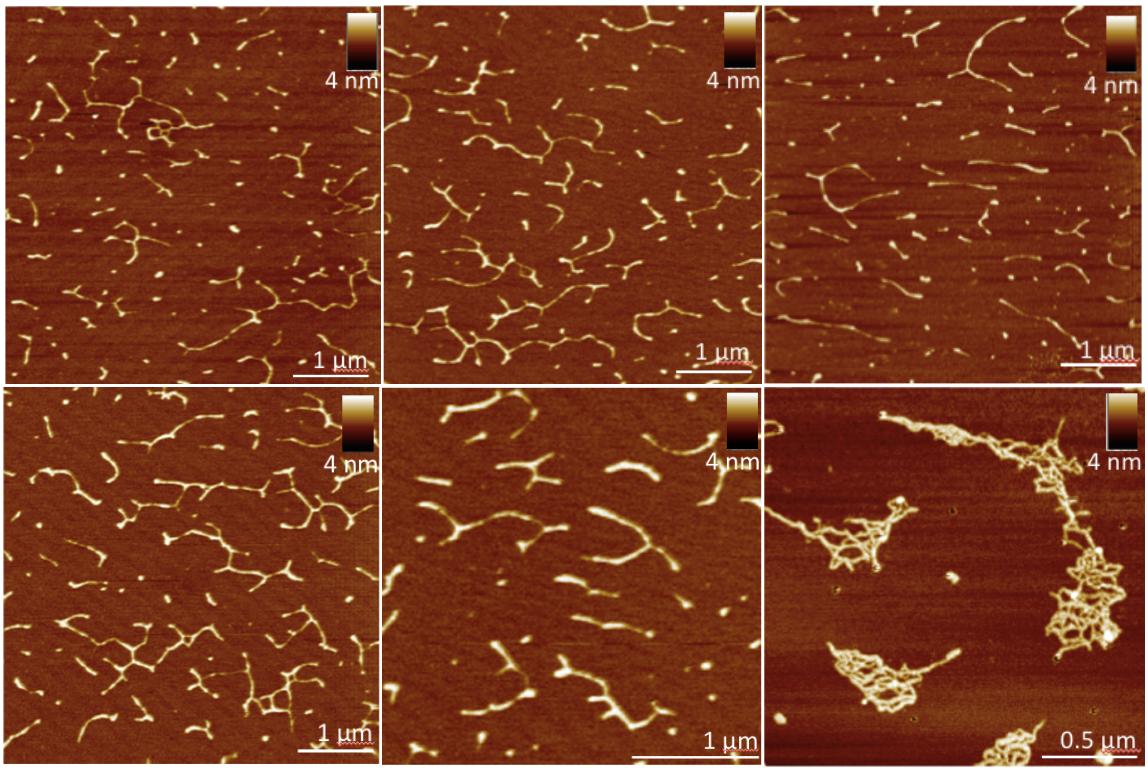


Figure 2.4. *AFM imaging of 4ng/μL 20 heat-cool cycle extDNA deposited on freshly cleaved mica with 0.5 mM* $MgCl_2$.

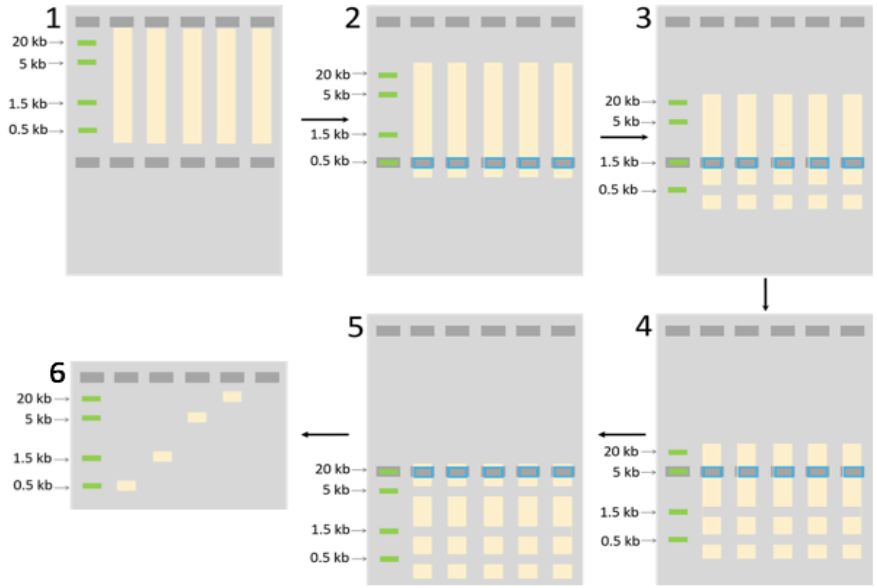
Firstly, it was clear that the extDNA sample used for the AFM imaging had grown to longer base pair lengths than that observed in the previous agarose gel, **Figure 2.2**, since the DNA on the agarose gel was too short to image well, < 300 bp. Images **A**-**E** revealed branching of the extDNA from several self-complementary strands of different lengths annealing together, in addition to the formation of folded DNA strands via a self-annealing mechanism. It was difficult to determine the exact lengths of the extDNA due to the branched and folded structures formed; however the height of the DNA was measured as 2 ± 0.3 nm. This height was higher than the height reported in the literature for dsDNA, 0.5 - 1 nm,[25] due to the self-annealing of the dsDNA and the structures containing more than two strands of DNA.

From the strands which could be measured (those which were not branched), the average length of extDNA was 530 ± 310 nm (1550 bp $\pm$ 900 bp), which was much longer than the average length depicted by the Image J analysis, **Figure 2.3**. The average length from the agarose gel may have been distorted by the formation of DNA structures which would inflate the number of "short" DNA strands. The average height was 1.4 ± 0.3 nm. Image **F** showed the formation of entangled DNA structures on the mica surface. The size of these structures varied depending on the base pair length of the DNA from which the structure arose. It was clear that the dsDNA was tightly entangled, and it was difficult to conclude if the structures were formed from one long strand of DNA self-annealing and entangling with itself, or if they were formed from multiple strands.

Figure 2.5. *Lonza™ gel electrophoresis 1.2% agarose two-tier pre-made gel cassette.*

Once the oligoseed had been extended, the wide distribution of DNA generated was narrowed by DNA length extraction to generate small aliquots of size-specific length DNA. Previous research used HPLC and gel extraction to produce DNA of specific lengths, however as discussed previously in **Section 2.1**, a more user-friendly and less

time-consuming method was desired. Lonza™ gel electrophoresis was a good candidate as it used a two-tier pre-made 1.2 % agarose gel cassette, **Figure 2.5**, thus removing the need to construct an agarose gel and the DNA could be extracted straight from the recovery wells. Furthermore, the requirement for gel slicing and purification was removed, alongside the need for lengthy HPLC runs where the peak had to be identified and collected. The Lonza™ gel method was termed "size recovery" and is outlined in **Scheme 2.3**.



Scheme 2.3. *Size Recovery using Lonza™ gel electrophoresis. Step 1: Extended DNA is inserted into the wells and run until slightly before the recovery wells. Steps 2-5: The user-defined length of DNA is extracted. Step 6: Purified size recovered DNA is run on a single tier 1 % agarose gel.*

The extDNA is inserted into the top tier of wells, alongside a relevant DNA ladder. A voltage of 250 V is applied, and the DNA migrates as the current passes through the gel; the shorter DNA lengths run faster than the longer length DNA, which creates a distribution of lengths. Before the desired length reaches the second tier of wells (known as the recovery wells), the current is stopped, and the recovery wells are emptied via micropipette then refilled with fresh Tris buffer (pH 7) solution. The current is restarted at the set voltage and the process repeated until all desired DNA lengths have been recovered. Size recovered products are purified with a commercially available PCR clean up kit which removes any dye accumulated during the running of the gel. The samples are then run on a standard 1 % agarose gel at 100 V in 0.5x TBE to determine the exact base pair lengths recovered.[17]

By using the size recovery method, the user is able to overcome the issue of not being able to obtain specific base pair length DNA from the Whitfield enzymatic extension method. Furthermore, it produces DNA lengths greater than 100 bp, which automated DNA synthesisers have trouble producing as the yield dramatically diminishes with the increase in oligo length.[26]

Size recovery was performed using $[AT]_n/[TA]_n$ extDNA, which had elongated to 1,500 bp during the Whitfield extension. Three size specific fractions were recovered at approximately 500, 700 and 1,000 bp in accordance with the GeneRuler 1 kb plus DNA ladder, **Figure 2.6**.

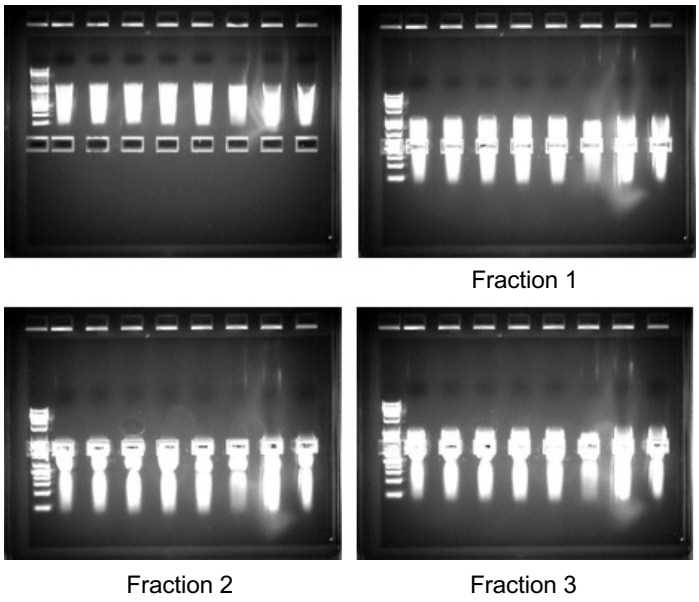


Figure 2.6. *Size recovery of extended $[AT]_n/[TA]_n$. Theoretical recovered base pair lengths were: fraction 1 = 500 bp, fraction 2 = 700 bp, fraction 3 = 1000 bp.*

It was evident from **Figure 2.6**, that small volumes of specific base pair length DNA were removed from the indents which were visible after the recovery wells. The gaps in the wide distribution indicated the length of the fraction removed by comparison to the ladder. Normally, the DNA extracted would have showed as black bands after the recovery wells, as seen in fraction 2, but the intensity of the gel was too bright in **Figure 2.6** to show clearly where the DNA had been removed from during size recovery. To determine the exact lengths of the collected fractions, standard agarose gel electrophoresis was performed and analysed by Image J, **Figure 2.7**.

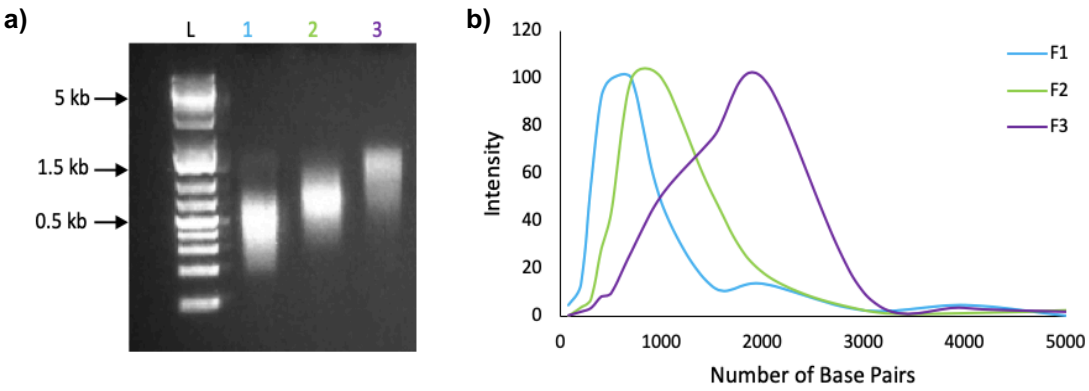


Figure 2.7.a) *Agarose gel of the size recovered fractions. Lanes 1-3: Fractions 1-3 and L = DNA ladder.* ***b)*** *Image J analysis of size recovered DNA fractions.*

Standard agarose gel electrophoresis was carried out using a 1 % agarose gel in 0.5x TBE and was run at 100 V for 1 hour. This ensured good separation of the chosen DNA ladder and the samples to be analysed. It was apparent that the fractions collected in lanes 1-3 did not form as tight a band as hoped and the recovered base pair lengths were still distributed across a wide range of lengths. Although the average base pair length of the fraction continued to increase and did not overlap with one another, the extremities of the band lengths bled into one another as indicated by the Image J analysis, **Figure 2.7b**. All three fractions contained a very large overlap with the neighbouring fractions which may have been due to the formation of concatemers after size recovery. The high intensity band of DNA in each of the three agarose gel lanes indicated that the DNA in this area was of lengths in good comparison to the DNA ladder as this band was very tight and formed a clear peak maximum in the Image J.

The faded areas on either side of the bright band suggested that there were other forms of $[AT]_n/[TA]_n$ DNA present, since the lengths of the faded distribution were out-with the DNA lengths recovered for that fraction. DNA of lengths below the bright band could be due to small self-annealed folded structures or small aggregates which were capable of flowing through the pores of the agarose. Those of base pair lengths greater than the bright band could have been from a very small amount of the next length of DNA entering the well and being extracted, however it was more likely that the faded distribution above the bright DNA band is a result of concatemeric DNA formations.

Aside from the potential presence of other DNA structures rather than only linear DNA, the tight DNA bands recovered, **Figure 2.7**, were much broader than aimed for during the size recovery process. This may have been due to the high running speed of the Lonza™

size recovery gel, 250 V, which would not have allowed the DNA to settle into the recovery wells. Due to the time taken to stop the voltage after identifying that the desired length was ready to be extracted, a greater amount of the unwanted longer length DNA had continued to enter the recovery wells. The speed of the gel was lowered to 150 V to minimise the length overlap between recovered DNA fractions. By reducing the migration speed of the DNA through the gel, the DNA would pass through the wells slower and greater accumulation of similar base pair length DNA in the recovery wells could be achieved.

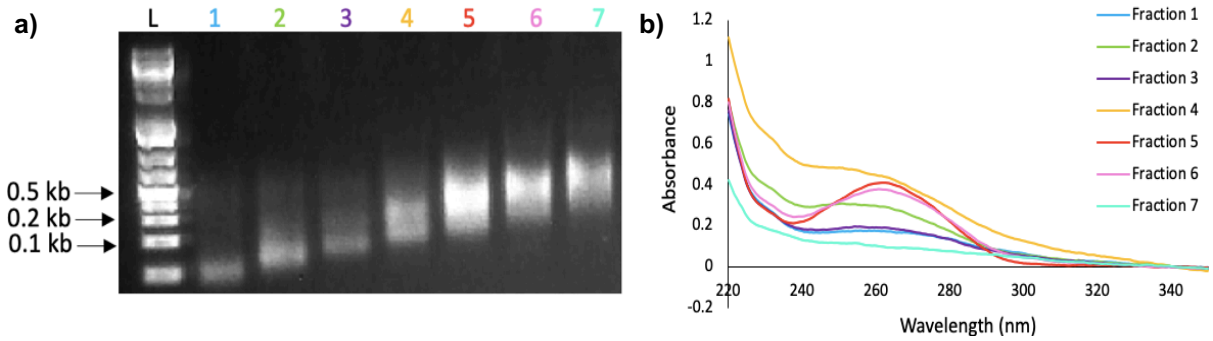


Figure 2.8. a) *1 % agarose gel of the size recovered fractions. Lanes 1-7: Fractions 1-7 and L = DNA ladder.* ***b)*** *UV-Vis spectroscopy, showing the 260 nm peak related to DNA for fractions 1-7*

Again, after size recovery and purification, a 1 % agarose gel was performed which showed the size of the recovered fractions, **Figure 2.8a**. It was evident that there were bright tight bands which were surrounded by a faded distribution of DNA lengths. The clear bands were the size specific DNA products, with the distribution thought to be made up of concatemers and DNA which had migrated too far and should not have been extracted. UV-Vis spectroscopy was used to further characterise the successful removal of the dsDNA and to obtain concentration values for each of the collected fractions. **Figure 2.8b**. In particular fractions 4-6 were more concentrated than the others recovered as a result of the Whitfield enzymatic extension which produced the average length of DNA within these sections.

Image J analysis, **Table 2.3**, showed that the base pair length ranges for all of the recovered fractions overlapped the neighbouring fractions. In the later fractions, this overlap was quite significant as fraction 7 had a base pair range of 375 - 1,500 bp and an average of 940 bp. The large range of base pairs reaffirmed that either the distribution was bleeding into one another during the size recovery process or DNA structures, such as concatemers, were formed post size recovery.

Table 2.3. *Concentration and calculated base pair lengths of size recovered fractions.*

Fraction No.	Abs @ 260 nm	Concentration (ng μL^{-1})	Base Pair Length – Mean (bp)	Base Pair Length – Range (bp)
1	0.174	8.7	100	20-175
2	0.296	14.8	190	125-250
3	0.192	9.6	250	100-400
4	0.446	22.3	350	175-525
5	0.404	20.2	450	225-675
6	0.378	18.9	550	275-825
7	0.302	15.1	940	375-1500

During the size recovery process, it was noticed that once the gel voltage was removed, the DNA still flowed through the gel and into the extraction wells for a period of time before it stopped. This resulted in the retrieval of DNA which was longer than the desired base pair length. Hence, to reduce the overlap in base pair lengths, the reduction in running voltage was kept constant and the DNA in the recovery wells was allowed to equilibrate for 10 seconds after the voltage was stopped before being removed. This ensured that all the DNA had stopped migrating through the gel before recovery of the desired length was undertaken, meaning that only the desired size was in the recovery wells, minimising the chance of a distribution of lengths being collected.

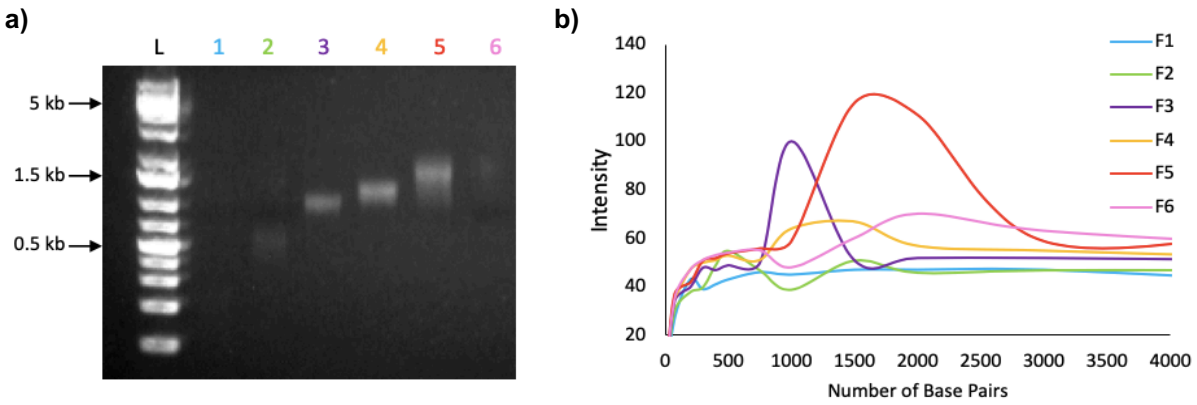


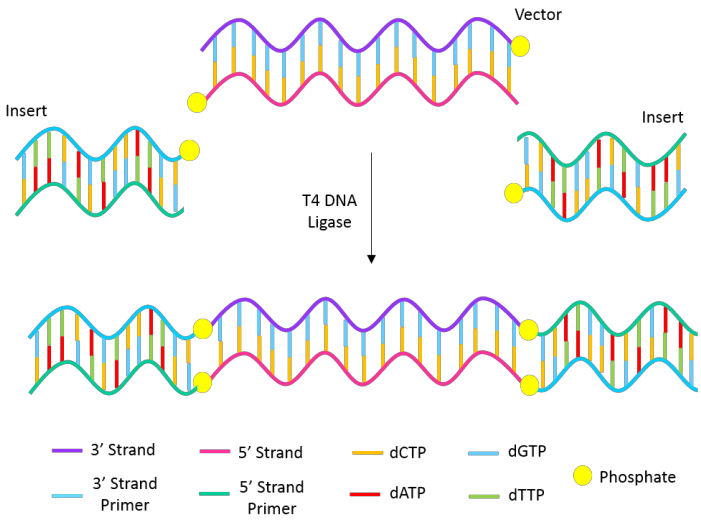
Figure 2.9. a) *1 % agarose gel of the size recovered fractions. Lanes 1-6: Fractions 1-6 and L = DNA ladder.* ***b)*** *Image J analysis of the recovered fractions.*

Narrower recovered DNA bands in the agarose gel were produced by allowing the gel to equilibrate after the voltage was stopped, **Figure 2.9a**, with no faded distribution evident in some of the fractions, emphasised in 3 and 4. However, fraction 5 showed a large faded

distribution, although it still contained a tighter band within the distribution above the band length of fraction 4. The Image J analysis, **Figure 2.9b**, showed sharp peaks for recovered fractions 3 and 4, which indicated the narrowness of the band. The streaking of the sample was more evident in fraction 5 as the base pair length range enveloped fraction 4 and 6, while it also bled a little into fraction 3. On the other hand, despite varying size recovered distribution lengths, it was clear that the peak maximums of the DNA fractions increased from fraction to fraction. This showed that the majority of the DNA was contained within the tight band observed on the gel and not in the faded ends of the band.

Furthermore, the intensity of fractions 3, 4 and 5 showed that these samples were more concentrated than the others and so, more DNA was removed during the size recovery process as the initial distribution was more concentrated at these base pair lengths. These fractions contained concentrations greater than 20 ng/μL, which were considered as high concentration values compared to other size recovered samples, < 10 ng/μL and fulfilled the aim of achieving aliquots of size specific DNA. However, they could not be used for future work as to be used for further experimentation or DNA assembly, the DNA would need to be recovered many times or processed synthetically to amplify the DNA to concentrations of 50 μM in 100 μL. Repeated size recovery would require many Lonza gels and experimental time, so the cost and effort would not be worth the increased recovered yield of size specific DNA. Therefore, amplification of the DNA fractions by synthetic PCR methods were undertaken.

Amplification of the size recovered DNA required a ligation step before PCR to prime the DNA for duplication using specifically designed primer sequences. Blunt end ligation was explored since the size-specific aliquot of DNA to be duplicated contained no "sticky ends"[27] and is considered a versatile method since the primer and target DNA have few sequence limitations. The process involved the joining of the primer to the exposed 5' phosphate on the target DNA by a DNA ligase,[28] **Scheme 2.4**.



Scheme 2.4. *Ligation Process. Insert DNA is attached to the vector DNA through the phosphate ester linkage using T4 DNA ligase.*[28,29]

T4 is the standard ligase used during ligation[29] as it is stable at room temperature and only requires incubation for 5 minutes at 25 °C for successful ligation. The designed primer sequence to be ligated to the target DNA was 5'-phosphate-TTAACGGCCAGATGGCCAGC-spacer C_3-3'. After the ligation process, the products were purified using a PCR clean up kit and characterised by agarose gel electrophoresis and UV-Vis spectroscopy, **Figure 2.10**.

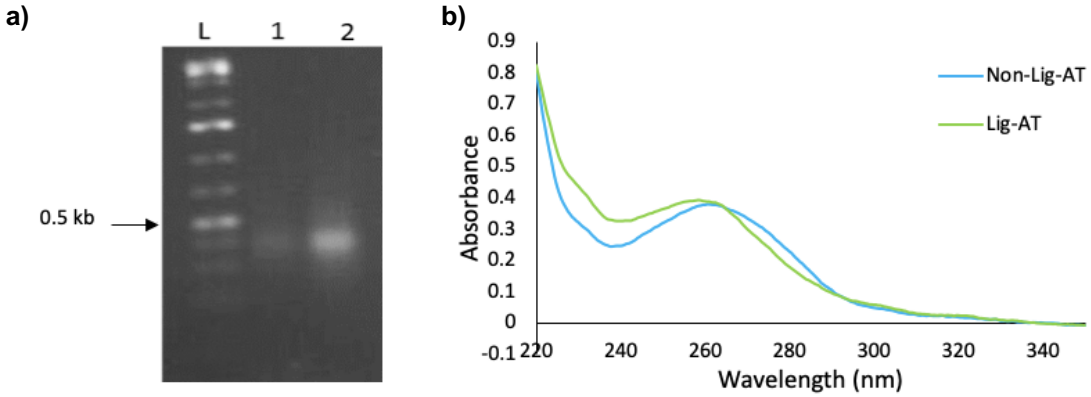


Figure 2.10. a) *Blunt end ligation with T4 DNA ligase and* $[AT]_n/[TA]_n$ *– fraction 6 (400 bp). Lane 1: non-ligated DNA, lane 2: ligated DNA and L = DNA ladder.* ***b)*** *UV-Vis spectroscopy of non-ligated and ligated samples.*

Figure 2.10 used 400 bp extDNA (fraction 6, **Figure 2.9**) for the blunt end ligation process. It was expected that a slight increase in the agarose gel length would be visualised due to the addition of 40 bp to the original sequence length, 20 bp on each 3' end of the target DNA. The extDNA band after ligation had shifted more towards 500 bp

on the agarose gel compared to un-ligated DNA. From the UV-Vis spectra, **Figure 2.10b**, it was evident that the 260 nm peak remained after ligation and at a similar absorbance when compared to the sample before ligation, the concentration was 18.9 ng/μL after primer addition compared to 19.6 ng/μL before. Not much variation in the UV-Vis spectra was expected after ligation, however it confirmed that no DNA was lost during the ligation and purification step.

As previously mentioned, PCR methods were investigated in order to achieve aliquots of size specific DNA which had concentrations of 50 μM in 100 μL. qPCR was the method of choice since the results of the amplification could be monitored in real-time.[30] qPCR can be performed after successful ligation by standard PCR protocol with the addition of a fluorescent dye. The fluorescent dye allows for real-time monitoring as the fluorescent signal is detected during the reaction by a fluorescence detector within the thermocycler.[31,32] This signal produces data which can be plotted as fluorescence intensity against number of cycles, **Figure 2.11**, where an increase in fluorescence indicates a positive result and no fluorescence a negative result. Hence, negative controls are used to confirm the amplification process.

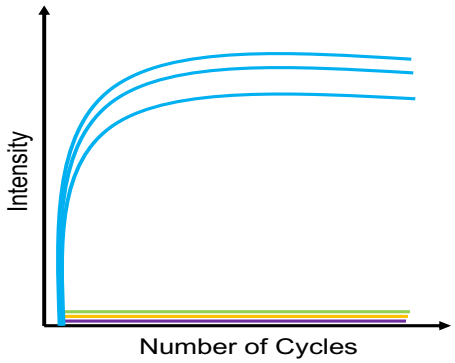


Figure 2.11. *qPCR amplification plot depicting the ideal amplification result with positive amplification of the target DNA (blue lines) and no amplification for the negative controls (green, yellow and purple lines).*[31,32]

In the qPCR method, samples were held at 95 °C for 120 seconds for full denaturation of the sample which was followed by 40 heat-cool cycles, performed in the presence of Pico Green (PG) dye to enable fluorescence monitoring. PG, **Figure 2.12**, is a fluorescent intercalator which displays an increase in fluorescence emission upon binding to dsDNA. It has a strong ability to interact with polymeric DNA and short duplexes. Dragan *et al.*[33,34] reported that the binding of PG is stabilised by the various interactions between the DNA

and PG where both the major and minor grooves of the DNA participate in complex formation. Wang *et al.*[35] supported this observation. Dragan *et al.*[33] described that intercalation of the complex into the major groove of the DNA occurs from the quinolinium group and the minor grooves are occupied by the dimethyl-amino propyl arms on the PG structure. Further stabilisation of the DNA-PG complex was reported by Dragan *et al.*[33] as occurring from the benzo-thiazole group which could form charge-charge interactions with the phosphate group. The fluorescence of PG increases 1000-fold on intercalation to dsDNA and is reported by Singer *et al.*[36] to be proportional to the quantity of DNA available, which therefore makes it a good DNA detector in qPCR.[37]

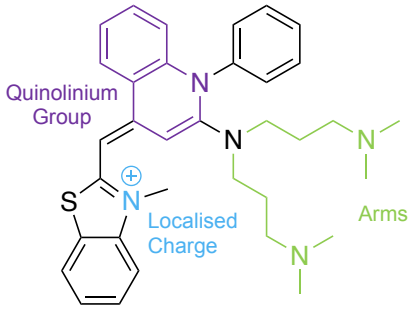


Figure 2.12. *Chemical structure of Pico Green depicting the different sections of the molecule involved in DNA binding; quinolinium group, dimethyl-aminopropyl arms and localised charge.*

The heat cool cycles consisted of standard qPCR conditions: 95 °C for 15 seconds, 60 °C for 15 seconds and 72 °C for 60 seconds. The annealing temperature (60 °C) took into consideration the melting temperature (Tm) of the primer duplex sequence, which is the temperature when one half of the DNA duplex dissociates to become single stranded and denotes the stability of the duplex.[38] After the blunt end ligation step, the target samples were prepared for qPCR alongside three controls which consisted of a positive extension control; non-ligated extDNA and two negative controls; one with no duplex and one with no polymerase.

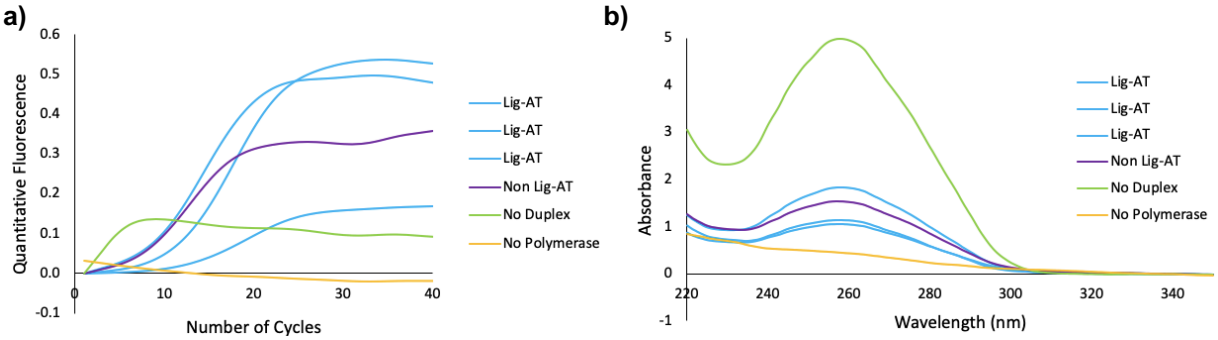


Figure 2.13. a) *qPCR normalised fluorescence, 3 ligated samples alongside three negative controls.* ***b)*** *UV-Vis spectroscopy for the qPCR samples and negative controls after purification.*

Figure 2.13a showed that the ligated ext$[AT]_n$ samples had amplified, alongside the no duplex negative control. This indicated that the ligated size-specific extDNA may not have only contained target DNA which had been amplified, as the behaviour of the primer could have also occurred in these samples. Thus, the fluorescence signal and concentration values would be elevated. UV-Vis spectroscopy showed the concentration of the amplified DNA after purification by a PCR clean up kit, **Figure 2.13b**. The ligated sample had increased from 18.9 to 66.7 ng/μL and so it was evident amplification of the DNA had taken place over the 40 cycles, although due to the high concentration of the no duplex negative control, 247.3 ng/μL, it could not be concluded that the target sample was not the only form or sequence of DNA present. It was determined that a non-ligated ext$[AT]_n$ negative control was not required as it continued to grow in length since the DNA was not ligated. The reaction mixture and PCR conditions for qPCR did not differ to the Whitfield extension method, and so the DNA would continue to extend. Therefore, growth would be observed in the qPCR fluorescence traces, as seen in **Figure 2.13**. For future qPCR experiments this control was not included.

To try to understand the results from the qPCR, a denaturing agarose gel was run after thermal cycling and purification, **Figure 2.14**. The denaturing agarose gel allowed for the identification of any small structures formed during the qPCR reaction or if the DNA elongated beyond its size recovered length through concatemeric structures. The concatemers are identified by separating the duplex DNA into two single strands and the individual strands migrate through the gel based on their own individual length rather than the duplex length. The dsDNA remains denatured by adding 3-(*N*-morpholino)propanesulfonic acid (MOPS) buffer and formaldehyde to the agarose solution during the setting of the gel. The same components are added to the DNA sample loaded into the wells and to the dsDNA ladder in order to ensure full denaturation of the

samples before loading. The gel was ran at 100 V for 1 hour, before post staining with ethidium bromide for viewing in the UV transilluminator.[39]

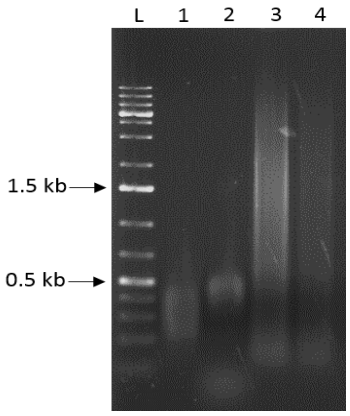
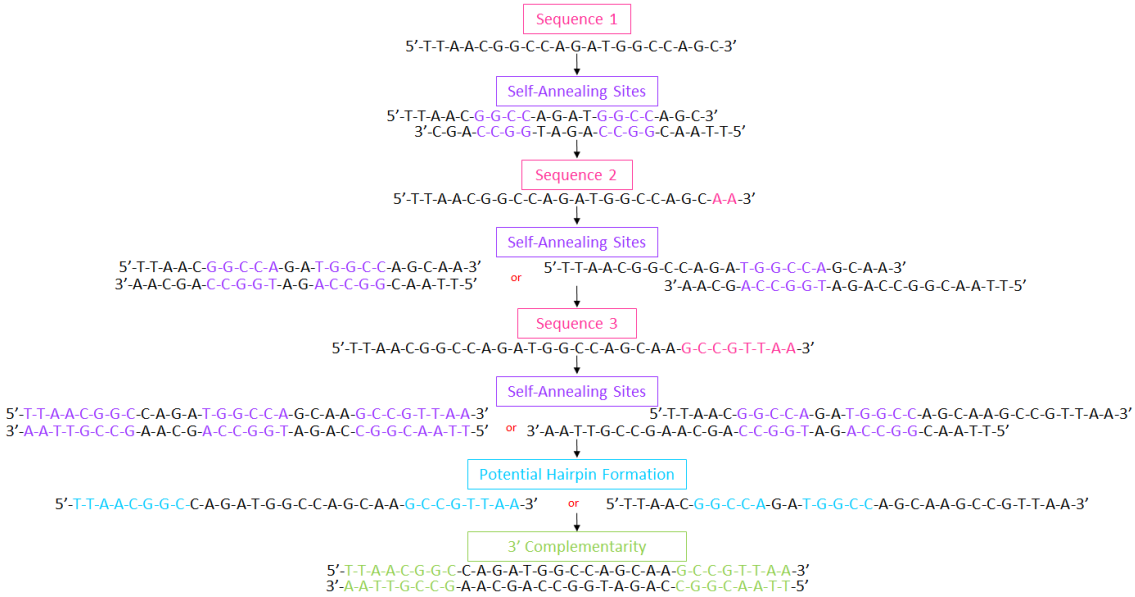


Figure 2.14. *Denaturing agarose gel of qPCR products. Lane 1: size recovered DNA, lane 2: size recovered and ligated DNA, lane 3: ligated and amplified DNA, lane 4: No duplex negative control (contains only primer DNA) and L= DNA ladder.*

Lane 1 and 2 showed the DNA as a small band of size recovered DNA at approximately 300 bp before qPCR, however; it was evident in lane 2 that after ligation a small band of DNA was visible below 75 bp. This was attributed to both small self-annealed structures which could migrate through the pores of the gel and concatemers of the primer DNA which had formed, including those which contained self-annealing and hair pin structures. These findings supported the no duplex negative control giving a positive response in the qPCR as the concatemers and self-annealed structures may have grown and encouraged intercalation of the Pico green dye, hence the fluorescence detection. This hypothesis was further illustrated by the result in lane 3, where a large distribution of DNA was present, 75 - 5,000 bp. The DNA in lane 4 was similar to the target DNA distribution in lane 3, 75 - 5,000 bp, and suggested that the target DNA was contaminated with extended primer DNA/concatemeric structures.

From the findings in **Figure 2.14**, theoretical information regarding the primer oligo was determined by the use of oligo calculator software developed by Northwestern University, known as OligoCalc[40] and would identify any issues with the primer. The calculations from OligoCalc highlighted that the primer sequence had 5 potential self-annealing sites during the first three cycles of qPCR and two potential hairpins, **Scheme 2.5**.



Scheme 2.5. *Potential primer sequence formations during qPCR cycling.*

As the results in **Scheme 2.5**, supported those found in **Figure 2.13** and **Figure 2.14**, it was decided that a new primer was required to address the issues with the current primer. The new sequence, 5' phosphate-TTAACGGCCAGATCGTCAGC-SpC_3-3', was found to have no self-annealing sites or potential hair pin formations associated with the sequence when entered into OligoCalc. Therefore, it should amplify the target sequence without primer side reactions occurring.

This new primer sequence was used on both bulk (extDNA not size recovered) and size recovered extDNA, to compare the ligation protocol of large distribution lengths to that of size recovered fractions, **Figure 2.15**.

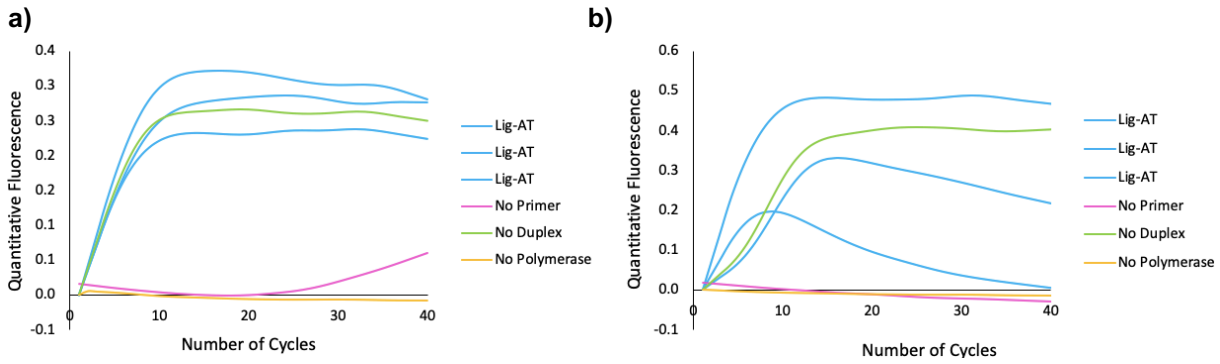


Figure 2.15. a) *qPCR with $[AT]_n/[TA]_n$ size recovered fraction* ***b)*** *qPCR with bulk $[AT]_n/[TA]_n$ containing a large distribution of lengths. In both qPCR data sets, 3 negative controls were included: No primer, no duplex and no polymerase.*

The qPCR traces of the size recovered extDNA, **Figure 2.15a**, showed a faster amplification when compared to the bulk extDNA, **Figure 2.15b**. The size recovered fractions had reached peak amplification by 5 cycles; however, the extDNA varied and two of samples did not reach peak amplification until 10 cycles. Therefore, the size recovered extDNA appeared to perform better than the bulk extDNA. Slow amplification in the bulk extDNA may have been caused by the large distribution of base pair lengths in the sample, as the DNA had more variations in length to form concatemers and self-annealed structures. This could have hindered amplification as the primer may not have been able to easily access the ligated primer on the target DNA to initiate amplification.

In both cases, three negative controls were run alongside the target DNA samples and again, it was apparent that the no duplex negative control had amplified when the trace should have been flat along the baseline. The trace for the no duplex sample was similar to that of the target extDNA and so, it was thought that the amplified target extDNA product trace was a result of the side-effect of the primer and not only amplifying the target extDNA fraction. The no primer negative control exhibited no fluorescence increase which indicated that there was successful blockage of the 3' end of the target extDNA sequence which inhibited the extension of the DNA by the Whitfield method.

To enable further analysis of the failed negative control by AFM, the qPCR products were purified using a NEB Monarch PCR and DNA clean up kit after thermal cycling. The no duplex negative control was fixed to the freshly cleaved mica surface at a concentration of 4 ng/μL in 1 mM $MgCl_2$, **Figure 2.16**.

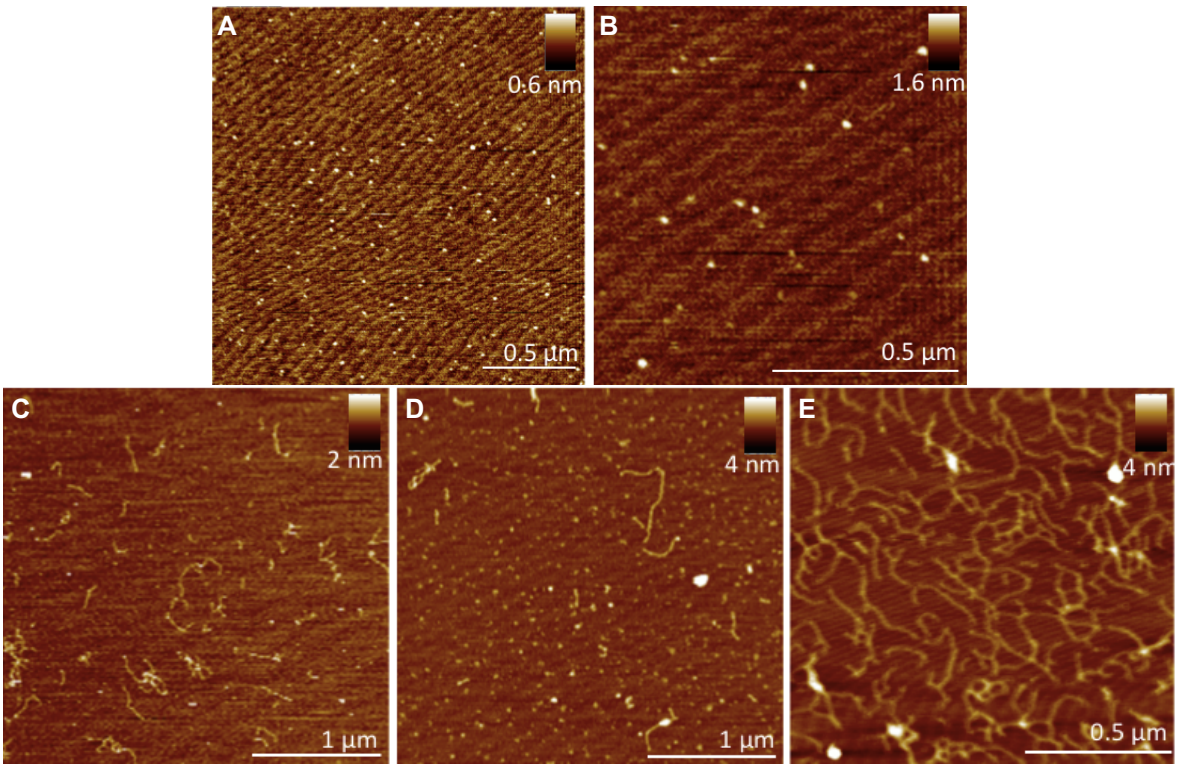


Figure 2.16. *AFM of the qPCR no duplex negative control sample containing only the single stranded primer DNA.* ***A-B:*** *before qPCR,* ***C-E:*** *after qPCR.*

Images A and B, **Figure 2.16**, showed that before qPCR the ssDNA of the primer was difficult to identify by AFM imaging as it appeared only as small dots of height 0.5 ± 0.1 nm, which would be expected for a sequence of 20 bases in length. In contrast, images **C**, **D** and **E** revealed how the single stranded primer had grown in base pair length during the qPCR thermal cycling and that the DNA was now double stranded due to the height, 1.0 ± 0.3 nm. The length of the primer DNA from **C** and **D** was an average of 200 ± 130 nm (590 ± 390 bp) which was an increase of approximately 570 bp from the starting oligo length.

Since the primer was deemed acceptable for use in qPCR from the results of the OligoCalc check; no self-annealing sites or potential hair pins, it was believed the primer was forming concatemers from base mis-matches which created dsDNA. Therefore, the enzyme could read the sequence and fill in any missing bases. In order to help determine whether the findings from the primer had any effects on the size recovered extDNA after qPCR, it was also imaged by AFM, **Figure 2.17**.

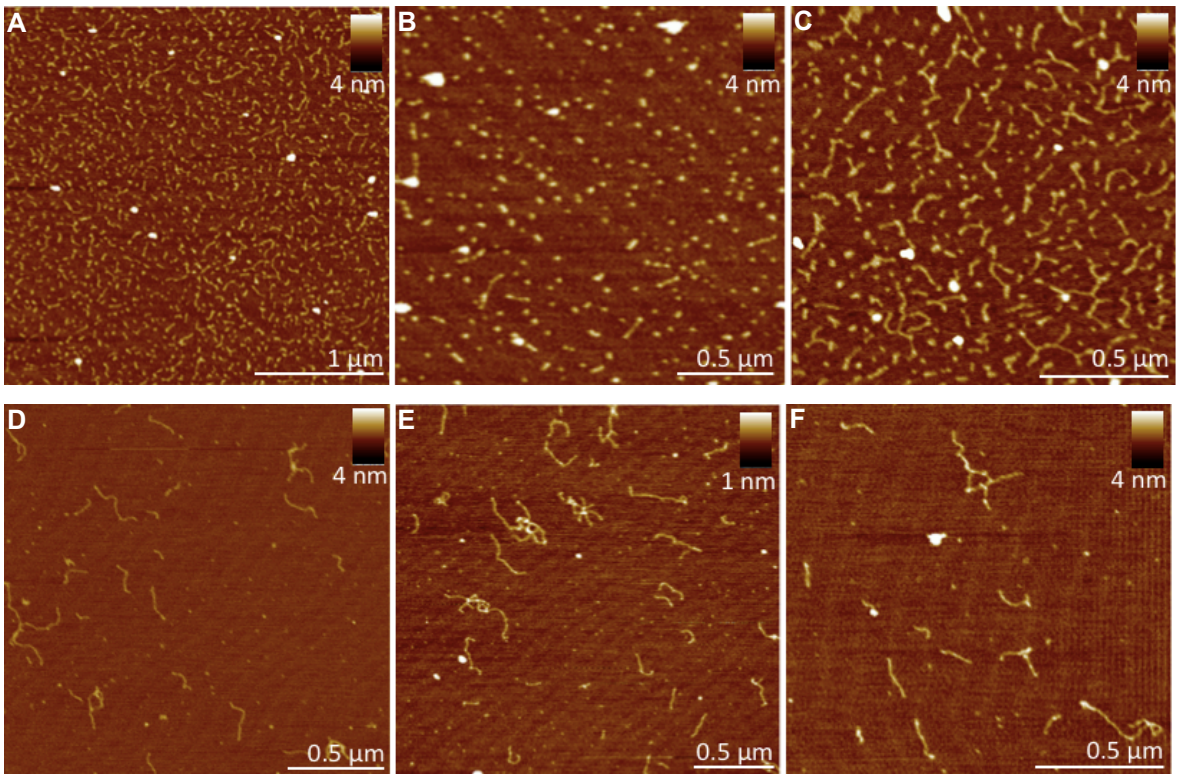


Figure 2.17. *AFM of qPCR ligated sample.* ***A-C:*** *before qPCR,* ***D-F:*** *after qPCR.*

AFM analysis of the size recovered ligated extDNA sample, **Figure 2.17**, indicated that before qPCR (**A**-**C**) the DNA was around 130 ± 20 nm (370 ± 60 bp). This sample before qPCR also contained a small range of lengths as could be witnessed from the variation between dots of DNA and stretched out lengths of DNA between the images. The measured height was 1.0 ± 0.2 nm, corresponding to the height of dsDNA. After qPCR (**D**-**F**), it was clear that the size recovered DNA had extended as there was a visible increase in the base pair length of the DNA due to the fact there were less DNA dots found in the image and the DNA was more aligned after molecular combing, which was easier to achieve with longer base pair length DNA. The DNA length had increased to 150 ± 80 nm (450 ± 250 bp), confirming elongation of the DNA from its original size recovered length, 370 bp. The height of the dsDNA was 0.8 ± 0.2 nm, which was within error of the original height.

The base pair length increases could be attributed to the primer behaviour in **Figure 2.16** but may have also been caused by Whitfield extension if the ligation procedure did not generate a 100 % yield. Although **Figure 2.15** showed the no primer negative control to respond as expected, the ligation was checked by inserting the same oligo primer that was attached to the target DNA.

The single stranded primer, containing the 5' phosphate and 3' C_3 spacer, was used in the reaction mixture as it would not bind to the ligated target DNA since the primer sequence was the same and not the primer complement. Therefore, with no primer to bind to the target DNA, the polymerase would have no dsDNA sequence to read and would not copy. If the target DNA were to form a minimum or maximum overlap to induce Whitfield extension, the DNA would not anneal properly due to the ligated primer sequence. Hence, it would block the polymerase from reading the sequence and the ligated samples would not give fluorescence traces, so ligation would have been successful.

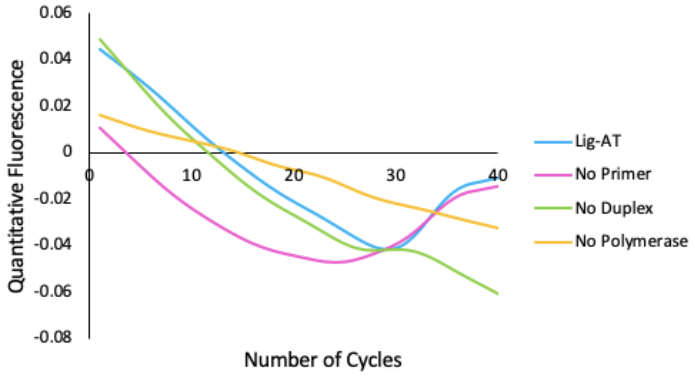


Figure 2.18. *qPCR with bulk $[AT]_n/[TA]_n$ using primer containing blocking regions which prevent amplification and enzymatic extension.*

Figure 2.18 agreed with the theoretical thought process of using ligation to block the 3' end of the DNA sequence so that it could not extend during qPCR amplification. None of the ligated target DNA samples or negative controls produced positive qPCR fluorescent signals and so, it was apparent that ligation of the size recovered extDNA was successful. However, from this data it was evident that the presence of both strands of the primer sequence in the qPCR reaction mixture was a contributing factor to the growth increase witnessed for the no duplex negative control. Furthermore, it indicated that the uncertainty in the qPCR detection was approximately ± 4 % as the signal measured changed over the course of the qPCR cycles.

Although ligation of the target DNA sequence was successful, it was apparent from previous results that the ligation protocol for DNA amplification was not yielding successful copies of purely target extDNA, a second protocol to incorporate a primer sequence was investigated; dA-tailing.[41] This protocol removed the need for a complicated primer which could potentially self-anneal, hairpin or form concatemers as a poly(A) tail was added to the 3' end of the target sequence. Therefore, for qPCR only a

poly(T) primer of 12 - 20 bp was required. The dA-tailing protocol was simple, only requiring incubation of the DNA at 37 °C for 1 hour with the NEBNext dA-tailing module.[42] In order to obtain suitable samples for amplification, the protocol was performed on the bulk extended $[AT]_n/[TA]_n$, as the concentration of the DNA to be dA-tailed was higher than if it was size recovered and so, reduced the risk of pipette error from having to use small volumes of reaction mixture.

Table 2.4. *Absorbance at 260 nm and concentration for $[AT]_n/[TA]_n$ before and after dA-tailing.*

Sample	Abs @ 260 nm	Concentration (ng/μL)
Before dA-tail	1.86	92.9
After dA-tail	2.12	106.2

It was difficult to characterise the dA-tail addition to the $[AT]_n/[TA]_n$ extended DNA as the number of bases added was not large enough to visualise a difference on an agarose gel and it did not add any new bases to observe the incorporation by HPLC digestion. UV-Vis spectroscopy was able to give some indication of successful dA-tail addition as a slight increase in concentration was observed after dA-tailing, **Table 2.4**. The length of the dA-tail added to the 3' end of the sequence was unknown as it was time and enzyme dependent, therefore, the longer the sample was incubated, the longer the dA-tail should have grown, with respect to the activity and half-life of the polymerase.

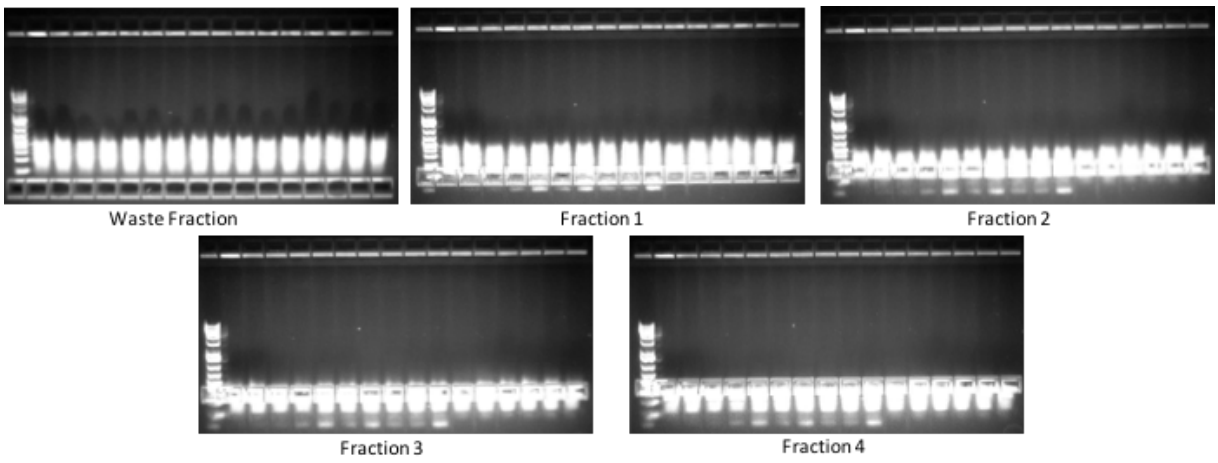


Figure 2.19. *Size recovery of extended and dA tailed $[AT]_n/[TA]_n$. Theoretical bp lengths are: fraction 1 = 75 bp, fraction 2 = 200 bp, fraction 3 = 200 - 300 bp and fraction 4 = 400 - 500 bp.*

Size recovery using the optimised method, as previously discussed, was performed on the dA-tailed bulk DNA to achieve size specific aliquots for amplification, **Figure 2.19**. In the

size recovery Lonza™ gel, four fractions were extracted with approximate base pair lengths of 75, 200, 300 and 400 - 500 bp. The tightness of the extended DNA distribution was evident and made extraction of the specific base pair lengths difficult as the individual lengths were hard to accurately determine.

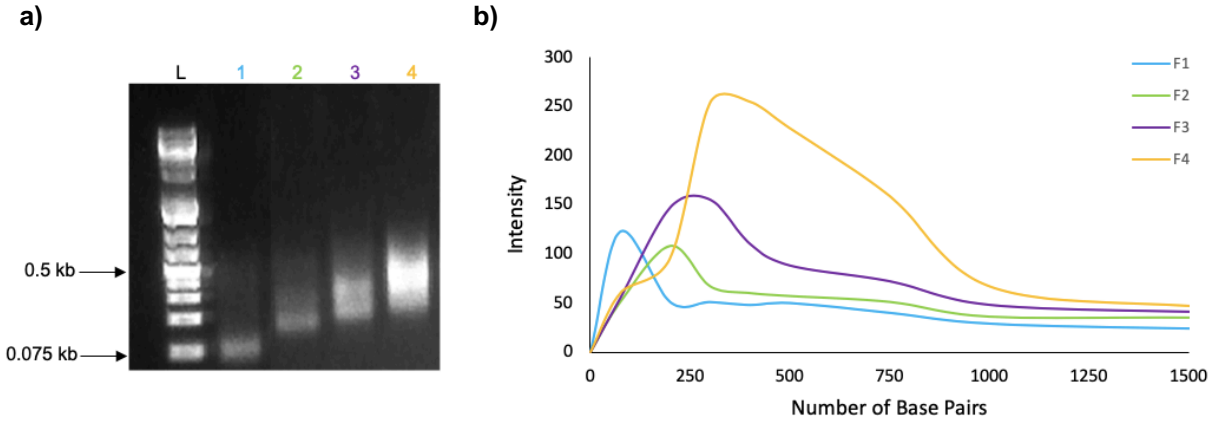


Figure 2.20. a) *Agarose gel of size recovered dA-tailed $[AT]_n/[TA]_n$ fractions. Lanes 1 - 4: Fractions 1- 4 and L = DNA ladder.* ***b)*** *Plot profile analysis of fractions 1 - 4 from size recovered agarose gel.*

After extraction and purification, the size recovered fractions were run on a 1 % agarose gel to accurately visualise the narrowed distribution of size selected recovered bases, **Figure 2.20a**. It was clear from the gel that narrow bands had been removed, with a slight DNA drag to longer base pair lengths. The faded distribution could have been due to the formation of concatemers after the size recovery process. On the other hand, it may be caused by the tight band of the initial extended DNA distribution making extraction of the desired base pair lengths more difficult. From the clear DNA bands on the agarose gel, the recovered DNA lengths were 75, 200, 200 - 400 and 250 - 500 bp.

In the Image J analysis, **Figure 2.20b**, fractions 1 - 3 showed a sharp peak maximum at the beginning of their traces, indicative of the tight band on the gel, but then showed a small bump after the peak before levelling out again which illustrated the DNA drag on the gel. Fraction 4 did not have as tight a peak maximum and showed as a much broader peak on the Image J graph, however; the graph displayed a step-wise increase in the average number of base pairs (peak maximum) recovered for each fraction, despite their ranges over-lapping, **Table 2.5**. The peak height of the fractions differed slightly to the lengths indicated from the gel as the Image J analysis was more accurate than reading the DNA ladder by eye.

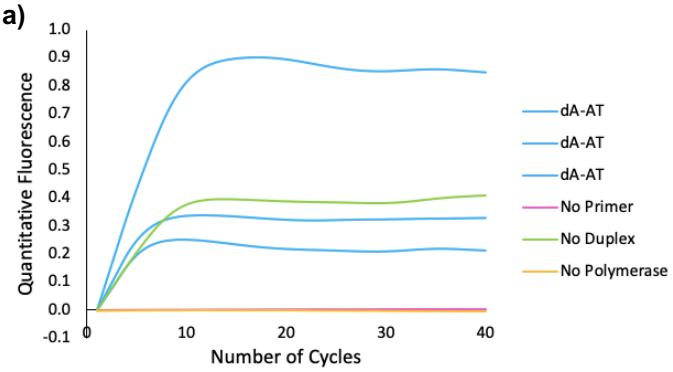
Table 2.5. *Absorbance at 260 nm of the size recovered dA-tailed DNA and the base pair lengths of each fraction.*

Fraction No.	Abs @ 260 nm	Concentration (ng μL^{-1})	Base Pair Length - Mean	Base Pair Length - Range
1	0.09	4.6	75	20-200
2	0.20	10.0	250	150-350
3	0.25	12.3	300	200-700
4	0.12	5.8	500	300-900

The UV-Vis spectroscopy data, **Table 2.5**, illustrated the importance of being able to amplify these DNA fractions as the concentrations after size recovery were relatively low and unusable for many applications so were required to be scaled up. For qPCR amplification of the dA-tailed size recovered fractions, the primer oligo was $[dT]_{20}$ as it was the complement to the dA-tail. There was no effect if the dA-tail was longer or shorter than the primer sequence. If longer then multiple primers would have bound and the polymerase would inset the corresponding bases into any spaces. If shorter, the primer would flank the end of the sequence, but the polymerase should be able to read the 5' to 3' direction of the sequence, in order to complete the copy.

Initially, a small study into the annealing time for the primer to interact with the dA-tail on the target sequence but not interact with itself during qPCR was investigated by using anneal times of 15 seconds and 5 seconds. These times were chosen because 15 seconds was the standard time used for all other qPCR experiments and 5 seconds was sufficient for the primer to find the dA-tail on the target sequence but not enough time to enable self-annealing through base mismatches. Temperatures and times of all other steps in the heat-cool cycles remained constant as per previous qPCR experiments.

a)



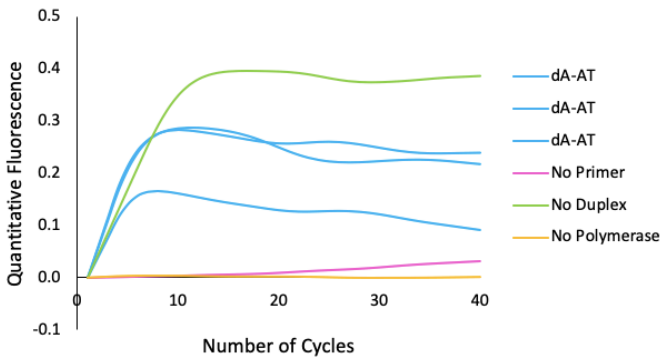
b)

Sample	Abs @ 260 nm	Concentration (ng/μL)
dA-AT	2.44	122.2
dA-AT	0.67	33.6
dA-AT	1.35	67.4
No Primer	0.20	9.7
No Duplex	1.04	52.2
No Polymerase	0.19	9.3

Figure 2.21. a) *qPCR using dA-$[AT]_n$/$[TA]_n$ fraction 2 (200 bp) with a 15 second anneal step.* ***b)*** *Absorbance and concentration at 260 nm after qPCR.*

Figure 2.21 used a 200 bp size recovered fraction of dA-tailed $[AT]_n$/$[TA]_n$ DNA with a 15 second anneal step during qPCR. Rather than affording no response and thus remaining as a flat line on the horizontal axis, the no duplex negative control gave an unexpected positive response. The other two negative controls gave no fluorescent response. From the concentrations obtained, **Figure 2.21b**, it suggested that the dA-AT sample was a mixture of target product and the primer forming structures during qPCR thermal cycling, as the concentration of the no duplex control, 52.2 ng/μL, was in close proximity to the average dA-AT sample concentration, 74.4 ng/μL.

a)



b)

Sample	Abs @ 260 nm	Concentration (ng/μL)
dA-AT	1.16	58.2
dA-AT	1.11	55.6
dA-AT	0.39	19.4
No Primer	0.11	5.9
No Duplex	0.90	44.8
No Polymerase	0.12	6.2

Figure 2.22. a) *qPCR using dA-$[AT]_n$/$[TA]_n$ fraction 3 (300 bp) with a 5 second anneal step.* ***b)*** *Absorbance and concentration at 260 nm after qPCR.*

The samples from the 5 second anneal time, **Figure 2.22a**, showed little difference in their fluorescent behaviour when compared to the 15 second anneal time results. Again, it was evident the dA-target DNA samples had increased in fluorescence, along with the no duplex negative control. This was also reflected in the concentration values recorded after purification of the samples, **Figure 2.22b**. The dA-AT sample concentration of 19.4 ng/μL was lower than the other dA-AT target samples, 58.2 and 55.6 ng/μL and could have been

due to pipette error when making the sample. The dA-target DNA concentrations did not differ much from the no duplex negative control, 44.8 ng/μL, so there was no advantage of using 5 second anneal time over a 15 second one. To obtain a better understanding of the qPCR results from using a dA-tail as the ligated primer, the samples were purified, and a 1 % agarose gel was run, **Figure 2.23**.

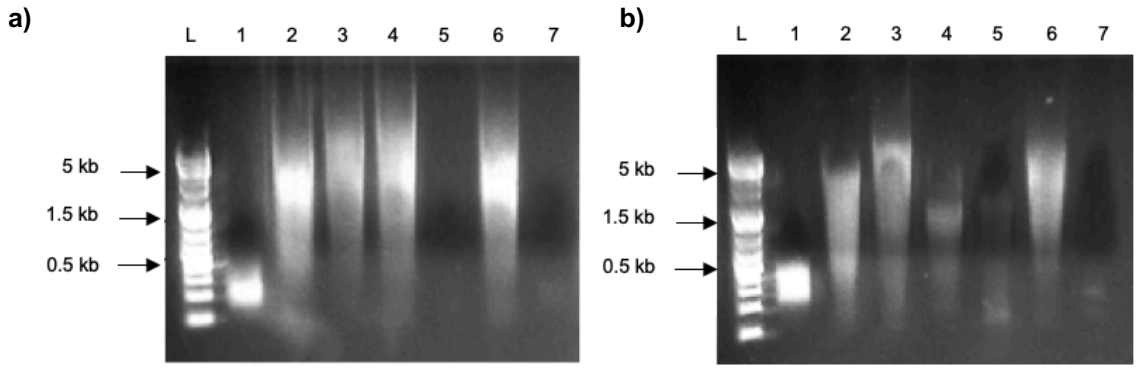


Figure 2.23. a) *Fraction 2 qPCR samples after purification. Lane 1: dA-$[AT]_n$ size recovered control, lanes 2 - 4: amplified dA-$[AT]_n$, lanes 5 - 7: negative controls, no primer, no duplex and no polymerase.* ***b)*** *Fraction 3 qPCR samples after purification. Lane 1: dA-$[AT]_n$ size recover control, lanes 2 - 4: amplified dA-$[AT]_n$, lane 5 - 7: negative controls; no primer, no duplex and no polymerase. L = DNA ladder.*

From lanes 2 - 4, **Figure 2.23a**, it was clear that the size specific dA-tailed target DNA had increased from 200 bp to an average of 2,000 bp and there was no clear band at 200 bp as per lane 1. However, it was also evident that the no duplex negative control had extended to the same extent as the target DNA samples, with the same band of DNA present at approximately 2,000 bp. This suggested that the target DNA samples had grown in length but was contaminated with the effects of the primer. The results discussed for the 15 second anneal, were echoed in the 5 second anneal agarose gel results, **Figure 2.23b**, which again confirmed that the anneal time had no effect on the qPCR amplification.

For the $[AT]_{10}/[TA]_{10}$ sequence it was found that amplification of the dA-tailed size specific DNA aliquots was difficult due to the effects of the primer. Solving the amplification issues was complicated as both the target sequence and the dA-tail contained dA bases. This meant the target sequence may have allowed T-T base mismatches and so, the primer sequence could have bound to the target and not just to the dA-tail on the end of the sequence which would have allowed for extension to occur.

A sequence which had no dA base present before dA-tailing would aid in the characterisation of the addition of the dA bases and may stop the primer annealing to the target through base mis-matches. Therefore, the effects of the primer on the target DNA would be removed from the qPCR results.

2.2.2 $[GC]_{10}/[CG]_{10}$ Oligo

$[GC]_{10}/[CG]_{10}$, a similar oligoseed to $[AT]_{10}/[TA]_{10}$, was chosen to undertake amplification studies without the presence of dA bases in the target sequence. Before amplification could be carried out the new oligoseed had to be extended to generate a large base pair length distribution, from which size-specific lengths could be extracted via size recovery.

Whitfield enzymatic extension was performed at 5, 10, 15 and 20 heat-cool cycles (95 °C for 30 seconds, 55 °C for 30 seconds, 72 °C for 120 seconds) with *Thermococcus gorgonarius* polymerase B (Pfu-Pol) Z3 exo$^-$,[18] to determine the optimum number of cycles to achieve the longest length of DNA, **Figure 2.24**.

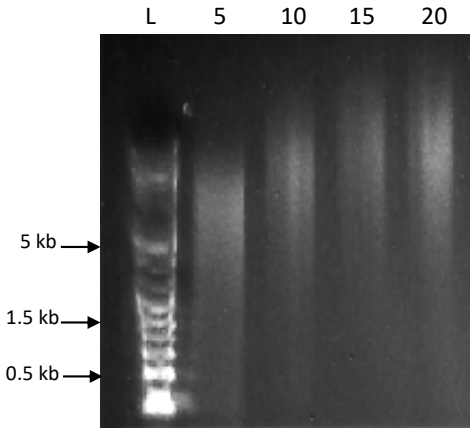


Figure 2.24. *Enzymatic extension with Tgo-Pol Z3 exo- and $[GC]_{10}/[CG]_{10}$ for 10, 20 and 30 cycles. Lanes 1 - 3 on the agarose gel show extension products after 10, 20 and 30 cycles, respectively and L = DNA ladder.*

The agarose gel showed that the DNA had grown to base pair lengths much longer than expected, especially after 5 cycles. The theoretical maximum length calculated for 5 cycles was 392 bp, which was much shorter than the length observed on the agarose gel. Furthermore, after 5 cycles, there was not a significant difference in the average base pair length between 10, 15 and 20 cycles. It was expected that a step-wise increase in the average length would be observed between the cycle lengths, especially as their theoretical base pair extension lengths were extremely different, **Table 2.6**.

Table 2.6. *Theoretical maximum and minimum DNA length after 5, 10, 15 and 20 enzymatic extension cycles.*

No. of Cycles	Theoretical Maximum (bp)	Theoretical Minimum (bp)
5	390	30
10	7080	40
15	17080	50
20	27080	60

Furthermore, the Image J analysis plot, generated from the agarose gel, did not help to analyse the base pair lengths as the DNA had grown to lengths longer than the correlating DNA ladder (maximum length 20,000 bp) and therefore, was not included. The unexpectedly long DNA base pair length was thought to have been generated due to the self-complementarity of the sequence and its ability to self-anneal or form concatemeric structures.

Table 2.7. *Absorbance and concentration at 260 nm for enzymatic extension with $[GC]_{10}/[CG]_{10}$ at 5, 10, 15 and 20 cycles.*

No. of Cycles	Abs @ 260 nm	Concentration (ng/μL)
5	0.99	49.4
10	1.86	92.9
15	2.21	108.9
20	2.70	135.0

Although the gel lengths of the extDNA did not produce the expected pattern on the agarose gel, the concentration/absorbance values did increase with respect to cycle number, **Table 2.7**. In order to try and understand how the $[GC]_n/[CG]_n$ DNA was behaving and to visualise the actual DNA strands, AFM of the 20-cycle product was performed at 4 ng/μL on freshly cleaved mica, **Figure 2.25**.

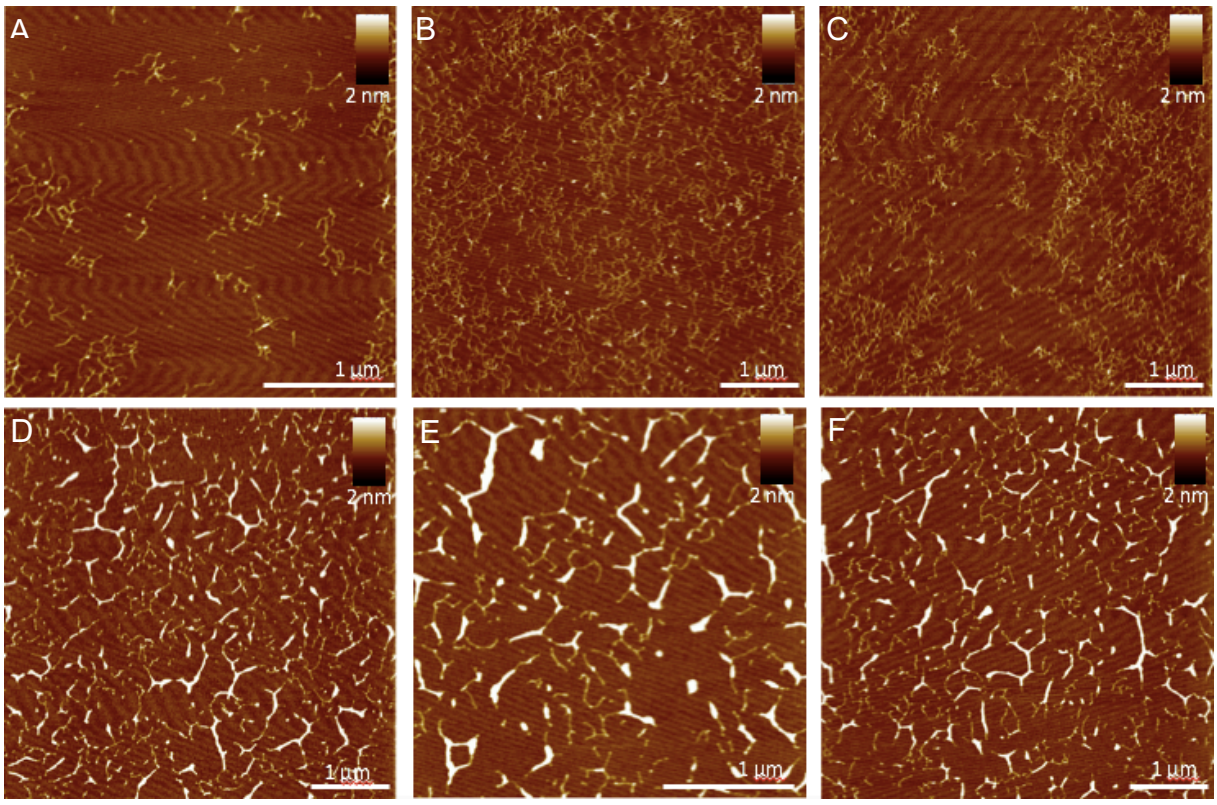


Figure 2.25. *AFM of $[GC]_n/[CG]_n$ after 20 cycles enzymatic extension with a 2 ng/µL sample on a freshly cleaved mica surface with 0.5 mM $MgCl_2$.*

From the long base pair length extDNA, different formations of DNA structures were observed in the AFM images. Images **A**-**C** showed clear entanglement of the long DNA strands on different areas of the mica surface. Image **A** denotes an area where the DNA was less concentrated than **B** and **C** but showed the folding formations in more detail due to a less cluttered image. Concentration differences across the sample was expected as the 4 ng/µL sample spread out over the mica surface during molecular combing, with any short DNA strands/structures typically settled towards the top of the mica and the longer strands/structures fixed near to the bottom of the mica. The high density of extDNA in **B** and **C** made it difficult to observe any measurable strands which were not entangled and therefore, could not obtain an accurate length measurement.

The height of the DNA in images **A**-**C** could still be obtained and had an average height of 0.7 ± 0.2 nm. Images **D**-**F** were taken from a different area on the mica, where the concentration of DNA was much less than that in images **A**-**C**, as was evident by the less dense population of DNA. The DNA in these images was less entangled due to the reduction in concentration in this area of the mica; however, it was evident that the DNA appeared to be much thicker in appearance than was expected for dsDNA, observable by

the brighter sections of the DNA. Folding and self-annealing was the cause of the thicker DNA sections due to the self-complementarity of the sequence. Again, it was difficult to determine the average length of the DNA, as many of the strands were involved in self-annealed structures, although the strands which were able to be measured gave an average length of 400 ± 220 nm (1180 ± 640 bp). The height of the DNA was measured for the folded and unfolded sections, 2.1 ± 0.4 nm and 0.7 ± 0.1 nm respectively.

From the DNA lengths and structural formations observed in **Figures 2.24** and **2.25**, it was evident the extension was not performing as expected, potentially due to the incorrect reaction conditions for this oligoseed. The magnesium concentration is important in PCR reactions for thermal stability of the polymerase.[43] Therefore, an investigation into the magnesium salt concentrations was undertaken to try and aid the polymerase to achieve DNA lengths more in line with the calculated theoretical lengths. The standard Whitfield extension PCR reaction mixture contained 1 mM $MgSO_4$, so concentrations of 2 mM and 4 mM were considered to increase polymerase stability.

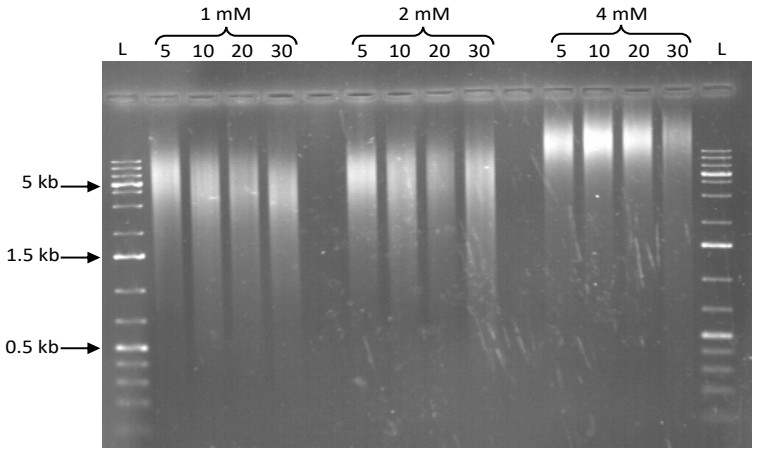


Figure 2.26. *Enzymatic extension with $[GC]_{10}/[CG]_{10}$ at varying concentrations of $MgSO_4$. 1mM $MgSO_4$ - lanes 1 - 4: 10 to 40 cycles, 2mM $MgSO_4$ - lanes 1 - 4: 10 to 40 cycles and 4mM $MgSO_4$ - lanes 1 - 4: 10 to 40 cycles. L= DNA ladder.*

It was clear from the agarose gel, **Figure 2.26**, that although the distribution did not display DNA lengths which correlated with the theoretical lengths, there were slight differences between the salt concentrations. The most obvious effect on the extDNA length was observed with 4 mM $MgSO_4$. For 1 and 2 mM $MgSO_4$, there was little difference between the base pair length distributions when using the different salt concentrations, hence there was no advantage in using 2 mM over 1 mM to obtain better extension parameters. The shift in DNA length was clear in the 4 mM samples as very tight, bright

bands of extDNA were identified at lengths of 20,000 bp. Since the DNA was longer than the ladder used for comparison, it was not possible to determine the exact length of the extDNA produced in the distribution.

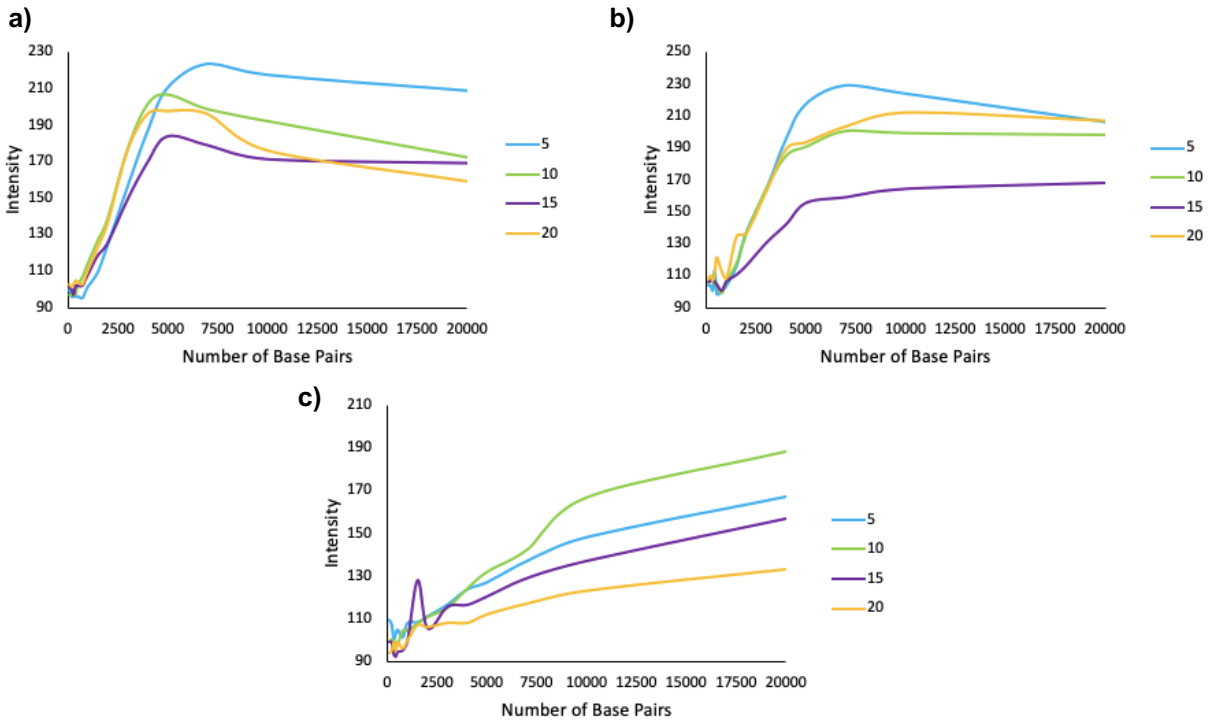


Figure 2.27. *Image J analysis of extDNA length after 5, 10, 15 and 20 Whitfield enzymatic extension heat-cool cycles with varying $MgSO_4$ concentration;* ***a)*** *1mM* ***b)*** *2mM and* ***c)*** *4mM.*

Image J analysis, **Figure 2.27**, further displayed the small variation in lengths produced by 1 mM and 2 mM $MgSO_4$ concentration. For all cycle lengths the peak maximums (average base pair lengths) for these samples were all at approximately 5,000 bp, with not much increase in length from the 5 to 20 cycles on the graphs, **Figure 2.27a** and **b**. **Figure 2.27c** showed that by increasing to 4 mM $MgSO_4$, the Image J traces shifted to longer base pair lengths, however no peak maximum was evident for these samples. This was due to the peak maximum being greater than the ladder length and therefore, was immeasurable by Image J, despite the DNA being clearly observed on the agarose gel.

Although changing the $MgSO_4$ salt concentration did not resolve the issue of extDNA which was much longer than the expected length after enzymatic extension, it did highlight how the extension length may be able to be controlled by salt concentration. So, rather than creating large distributions with every extension, tight bands of DNA may be obtained instead. At this time, this idea was not explored and the focus of the investigation into the PCR reaction conditions was shifted to the heat-cool cycle lengths and their effect on the extDNA length. It was thought that by reducing the number of cycles, smaller

distributions of DNA may be obtained. Although the theoretical base pair lengths did not correlate well to 5, 10, 15 and 20 heat-cool cycles, it was decided that less heat-cool cycles would give the oligoseed less chance to self-anneal or form concatemeric DNA, so may behave more like its theoretical length, **Table 2.8**.

Table 2.8. *Theoretical maximum and minimum DNA length after 1-10 heat cool cycles during the Whitfield enzymatic extension.*

No. of Cycles	Theoretical Maximum (bp)	Theoretical Minimum (bp)
1	32	22
2	56	24
3	104	26
4	200	28
5	390	30
6	780	32
7	1540	34
8	3080	36
9	5080	38
10	7080	40

The theoretical maximum length for 1 to 10 cycles ranged from 32 to 7,080 bp. However, these lengths were not reflected in the agarose gel, as even for one PCR heat-cool cycle length, very long lengths of extDNA were observed, **Figure 2.28**.

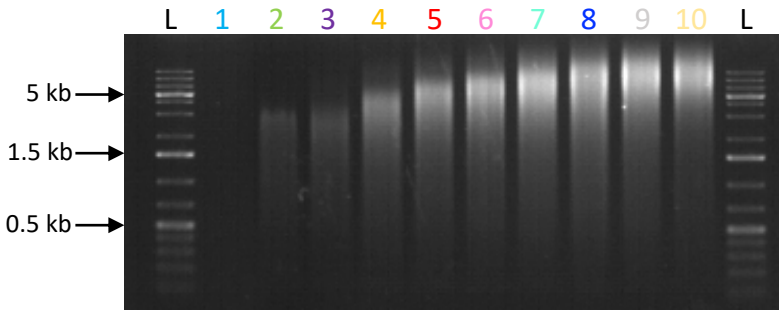


Figure 2.28. *Enzymatic extension with Tgo-Pol Z3 exo- and $[GC]_{10}/[CG]_{10}$ for 1 to 10 cycles. Agarose gel: lanes 1-10: extension products after 1, 2, 3, 4, 5, 6, 7, 8, 9 and 10 cycles, respectively and L = DNA ladder.*

Even though none of the heat-cool cycle lengths displayed extDNA lengths correlating to their maximum theoretical base pair length, an increase in step-wise growth was evident from cycle to cycle. This growth increment was clear until cycle 8 and thereafter, no

increase in length was observed. The difference in average base pair length (peak maximum) was shown in more detail by the Image J graph, **Figure 2.29**.

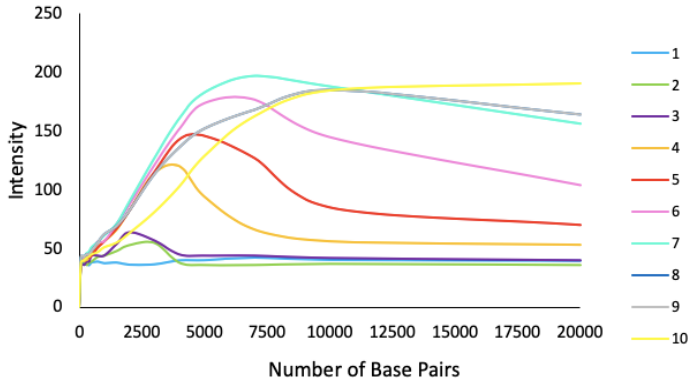


Figure 2.29. *Plot profile of enzymatic extension growth over 1-10 heat cool cycles with* $[GC]_{10}/[CG]_{10}$.

A minor increase in growth up to seven cycles was identified by the shift in peak maximum to longer base pair lengths. Thereafter, the number of base pairs started to plateau. For cycles four onwards, the trace did not decrease back to the baseline as the extDNA had extended further than the base pair length maximum of the ladder, 20,000 bp. Any DNA longer than 20,000 bp could not be plotted against the ladder lengths to produce an accurate peak maximum.

Instead of trying to obtain smaller distributions of $[GC]_n/[CG]_n$ by changing the reaction conditions or number of heat-cool cycles used for the Whitfield enzymatic extension, the standard reaction conditions for enzymatic extension was employed and size recovery was used to obtain DNA aliquots in the range 200-1,500 bp. A Lonza™ double tier 1.2 % agarose gel was used, and six DNA fractions were removed at approximate lengths of 100, 200, 400, 500, 1,000 and 1,500 bp, **Figure 2.30**.

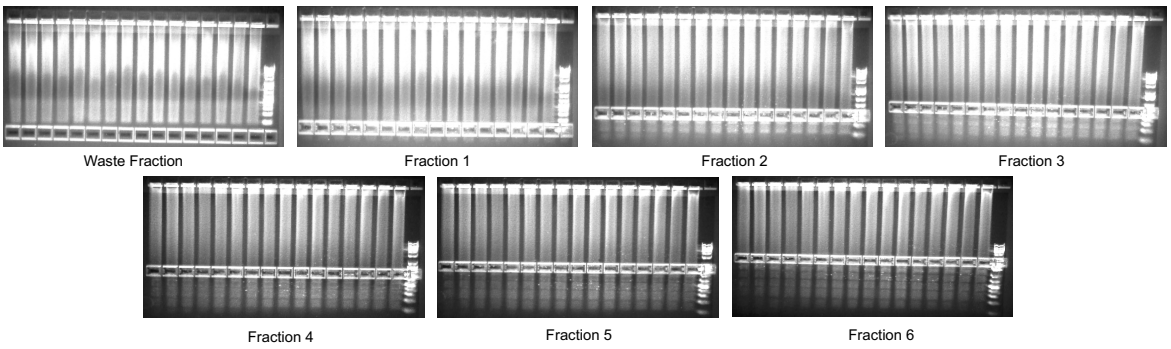


Figure 2.30. *Size recovery of extended [GC]n/[CG]n. Theoretical bp lengths are: fraction 1 = 75 - 100 bp, fraction 2 = 100-200 bp, fraction 3 = 300-400 bp, fraction 4 = 500 bp, fraction 5 =750 - 1000 bp and fraction 6 = 1500 bp.*

After size recovery and purification, a 1 % agarose gel was run to determine the actual DNA lengths of the size recovered fractions, **Figure 2.30a**. Condensed band lengths of DNA were clearly visible on the gel due to the use of previously established optimum size recovery conditions. Clear step-wise increments in length were obvious, which signified successful extraction of the desired base pair lengths as they did not bleed and overlap the other fractions. Furthermore, the bands of DNA were not intense which suggested the concentration of the size specific recovered fractions were low.

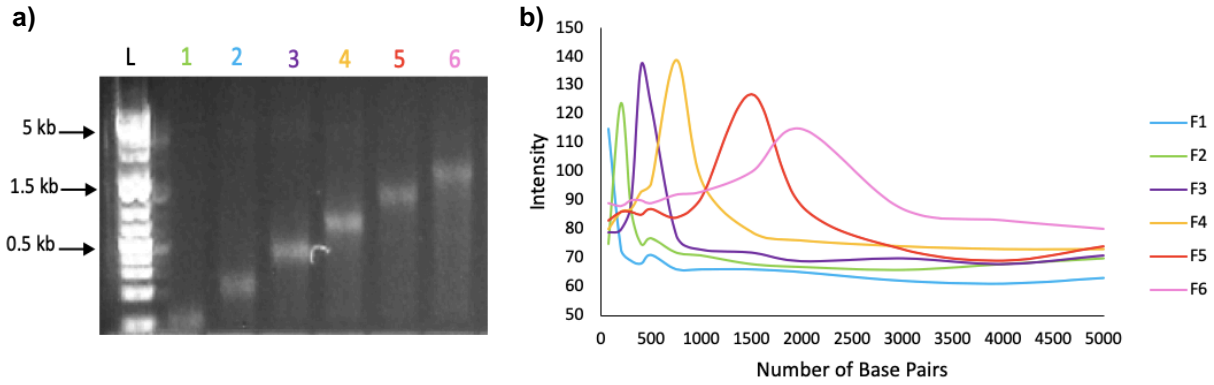


Figure 2.31. a) *Agarose gel of the size recovered fractions. Lanes 1-6: Fractions 1-6 and L = DNA ladder.* ***b)*** *Length analysis by Image J for fractions 1-6.*

The Image J analysis, **Figure 2.31b**, presented that on moving from short base pair length to long base pair length on the horizontal axis of the graph, the peak maximum for the average base pair length continued to increase and shift to the right from fraction to fraction. This indicated that there was successful recovery of short length ranges of extended $[GC]_n/[CG]_n$ DNA. There was still some length overlap in the size recovered fractions as the ranges of DNA collected, **Table 2.9**, sometimes contained lengths of DNA which belonged to the fraction before, although this overlap was minimal. Analysis by

UV-Vis spectroscopy showed that the concentrations of the size recovered fractions were very low since all samples were less than 10 ng/μL, so to obtain usable concentrations of DNA, amplification of the size specific DNA aliquots was required.

Table 2.9. *Concentration from UV-Visible spectroscopy and calculated bp lengths of size recovered fractions from plot profiling.*

Fraction No.	Abs @ 260 nm	Concentration (ng μL^{-1})	Base Pair Length - Mean (bp)	Base Pair Length - Range (bp)
1	0.11	5.69	75	50-150
2	0.08	4.10	200	75-250
3	0.12	5.78	400	310-510
4	0.13	6.25	750	550-900
5	0.10	4.77	1100	840-1650
6	0.05	2.48	1900	1150-2700

Due to the findings from the amplification of $[AT]_{10}/[TA]_{10}$ DNA in **Section 2.2.1**, it was decided that the dA-tailing technique should be used with $[GC]_{10}/[CG]_{10}$ DNA, since it contained no dA or dT bases which the poly(T) primer could compete with. Unlike with the $[AT]_n$ DNA, the size specific aliquots were dA-tailed individually rather than in bulk, to see if there was any observable difference in the agarose gel after poly(A) tail addition, **Figure 2.32**.

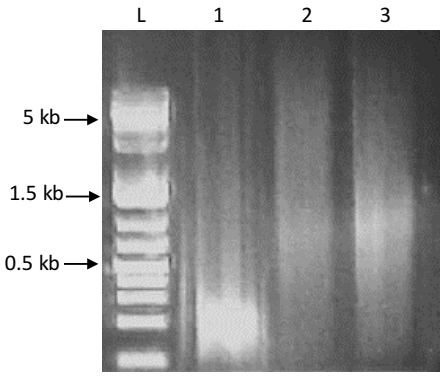


Figure 2.32. a) *1 % agarose gel. Lane 1: [GC]n/[CG]n extended DNA, lane 2-3: dA-tailed [GC]n/[CG]n. L = DNA ladder.*

The agarose gel was deemed to be poor in quality as the DNA ladder was not well resolved and the samples showed large streaks across the length of the gel which, in the case of the

ladder and lane 1, was not due to processivity of the samples before analysis. Despite this observation, the difference in length between the size recovered sample and the dA-tailed size recovered sample was evident. The gel image appeared to show that the dA-tailed DNA had grown slightly longer in length with a streaky distribution. The streaking of the non-dA-tailed sample, lane 1, lead to the conclusion that the dA-tailed extDNA had not actually grown in length to the full distribution indicated on the gel, since the streaking was clear in both samples. It suggested that the agarose gel was poorly executed as previously the size recovered sample was a tight DNA band, **Figure 2.31**.

After characterisation of the dA-tail by agarose gel, further confirmation of successful addition was obtained by HPLC. This allowed for identification of the individual bases present in the DNA sample, as each base is correlated to their own retention time as they flow off of the reverse phase column. Before injection into the HPLC, the dA-tailed extDNA was digested with snake venom phosphodiesterase at 37 °C overnight. The snake venom breaks down the dsDNA into single bases by cleavage at the phosphate, thus the bases can be individually detected, and retention times compared to the dNTP standards.

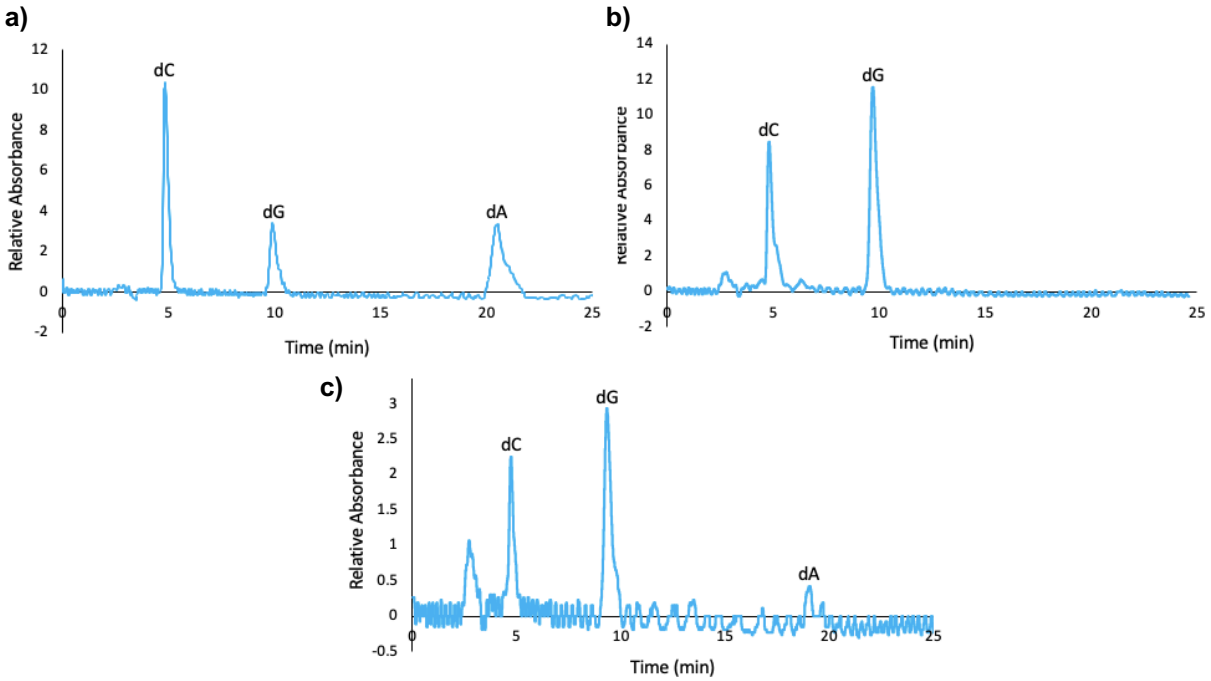


Figure 2.33. *HPLC traces using 2.5% acetonitrile (pH 6.5).* ***a)*** *dNTP standards with dCTP, dGTP and dATP.* ***b)*** *Digested* $[GC]_{10}/[CG]_{10}$ *oligo seed.* ***c)*** *Digested extended and dA-tailed* $[GC]_n/[CG]_n$.

Firstly, standard dNTPs of bases dC, dG and dA were examined to determine their retention times, **Figure 2.33a**. Digested extended $[GC]_n/[CG]_n$ was run as a control, which showed there was no dA base present in the size recovered sample before it was dA-tailed,

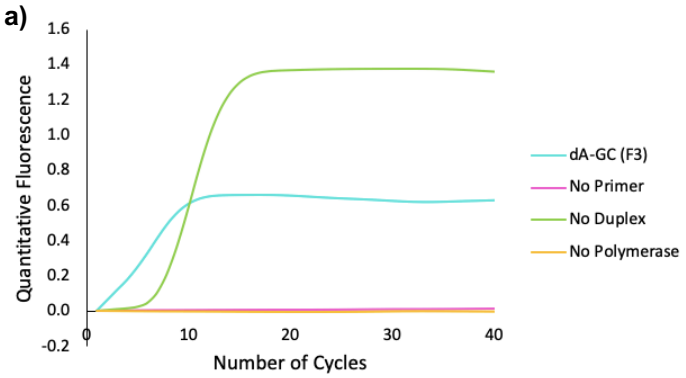
Figure 2.33b. Finally, digested extended dA-$[GC]_n/[CG]_n$, **Figure 2.33c,** was injected into the HPLC and displayed peaks corresponding to dC and dG, along with a small peak for dA. The peak for the dA base in the dA-tailed sample appeared very small in height on the HPLC trace, however this was expected as the concentration of dA base was minimal in comparison to the dC and dG concentration, especially if there was low addition of dA base during the tailing process.

Table 2.10. *HPLC retention times for dA-tailed $[GC]_{10}/[CG]_{10}$ extended dsDNA.*

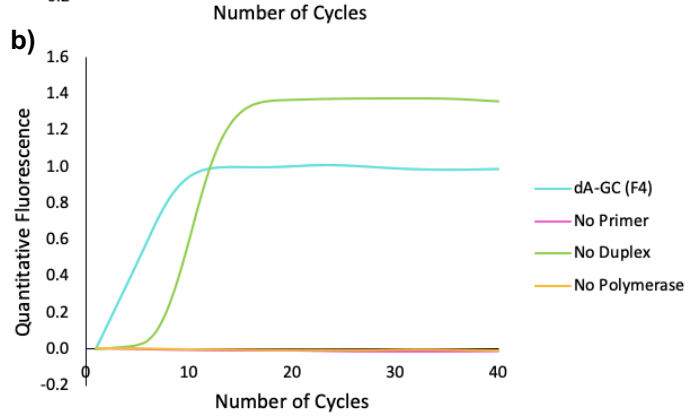
Sample	Retention Time (mins)			
	dC	dG	dT	dA
dNTP standards	4.9	9.9	-	20.6
$[GC]_{10}/[CG]_{10}$ oligo seed	4.8	9.7	-	-
Extended and dA tailed $[GC/[CG]_n$	4.8	9.3	-	19.1

The retention times of the $[GC]_n$ control and the dA-tailed extDNA sample were compared with the dNTP standards, **Table 2.10**. It was found that the retention times for dC and dG were very similar in all three samples at approximately 5 and 10 minutes, respectively. There was a small mis-alignment with the dA base between the dNTP standards and the dA-tailed extDNA sample, however there would be no other bases found in the vicinity of this area of the graph and therefore, the small discrepancy could be disregarded.

The dA-tailing of the $[GC]_n/[CG]_n$ extDNA was characterised to show that the dA-tail was present on the 3' end of the DNA strand, a detail which could not be fully concluded for the $[AT]_n/[TA]_n$ sequence. Therefore, $[GC]_n/[CG]_n$ samples were good candidates for amplification by qPCR. The qPCR conditions and number of heat cool cycles were unchanged from the amplification of $[AT]_{10}/[TA]_{10}$.



Sample	Abs @ 260 nm	Concentration (ng/μL)
dA-GC (F3)	0.58	28.9
No Primer	0.13	6.4
No Duplex	1.36	68.2
No Polymerase	0.10	5.0

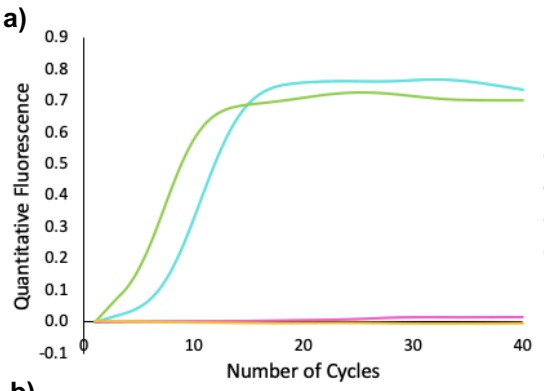


Sample	Abs @ 260 nm	Concentration (ng/μL)
dA-GC (F4)	1.32	66.0
No Primer	0.11	5.7
No Duplex	1.36	68.2
No Polymerase	0.10	5.0

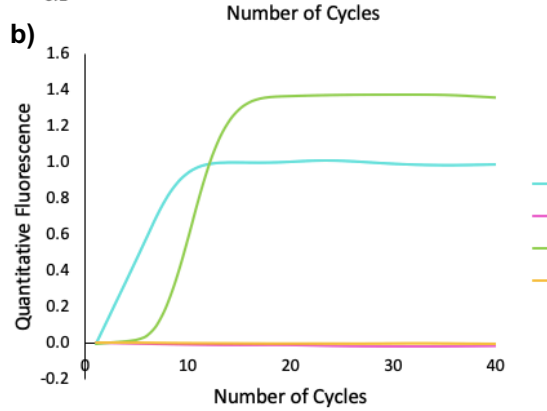
Figure 2.34. *qPCR with a 5 second anneal time of size recovered dA-tailed $[GC]_n$/[CG]n fractions and concentration of samples and their negative controls after amplification.* ***a)*** *Fraction 3 (400 bp)* ***b)*** *Fraction 4 (750 bp).*

As discussed for the amplification of the $[AT]_{10}/[TA]_{10}$ DNA in **Section 2.2.1**, the annealing time of the primer to the dA-tail was investigated. A 5 second anneal time was tested by using two size recovered fractions, fraction 3 (400 bp) and fraction 4 (750 bp) for qPCR, **Figure 2.34a** and **b** respectively. It was evident that both fractions produced the same fluorescent responses, where the dA-GC extDNA increased in fluorescence and so too did the no duplex negative control. This meant that despite no target DNA for the poly(T) primer to bind to, the primer was acting promiscuously, by forming dsDNA structures. This allowed for the PG dye to intercalate and fluoresce; hence it was detected by the qPCR. This was further confirmed by the absorbance data collected at 260 nm, **Figure 2.34**, as the concentration of the no duplex negative control was higher than that of both target DNA samples. This suggested that the amplification portrayed in the qPCR fluorescence traces for the target DNA samples actually belonged to the no duplex negative control and so, the target sample was not amplified to the intended sequence.

A 15 second anneal time was also tested to see if this would solve the primer issue by ensuring the primer had enough time to anneal to the dA-tail on the target DNA and form its most stable duplex form, **Figure 2.35**.



Sample	Abs @ 260 nm	Concentration (ng/µL)
dA-GC (F3)	0.32	15.9
No Primer	0.13	6.7
No Duplex	0.97	48.3
No Polymerase	0.26	13.2



Sample	Abs @ 260 nm	Concentration (ng/µL)
dA-GC (F4)	0.92	46.1
No Primer	0.48	23.9
No Duplex	0.97	48.3
No Polymerase	0.41	20.5

Figure 2.35. *qPCR with a 15 second anneal time of size recovered dA-tailed [GC]n/[CG]n fractions and concentration of samples and their negative controls after amplification.* ***a)*** *Fraction 3 (400 bp)* ***b)*** *Fraction 4 (750 bp).*

The same size recovered fractions of DNA were used in **Figure 2.35** as were used for the experiment in **Figure 2.35** and allowed for ease of comparison between the two annealing times. Unfortunately, no difference was observed in the 15 second anneal time fluorescence compared to the 5 second anneal time as the dA-GC sample increased in fluorescence as well as the no duplex negative control. The absorbance data at 260 nm showed an increase in fluorescence for the dA-GC sample which was more than likely related to the increase seen in the no duplex negative control. This was because the concentration values were either similar or higher for the no duplex negative control than for the dA-GC sample. Therefore, the dA-GC amplification appeared to contain other species of DNA than just the target DNA, so the DNA present in the sample was more attributed to the effect of the primer rather than amplification of the target sample.

In order to investigate the behaviour of the no duplex negative control product, a 1.2 % Lonza™ agarose gel was run with all purified products to determine how each sample responded to the qPCR amplification, **Figure 2.36**.

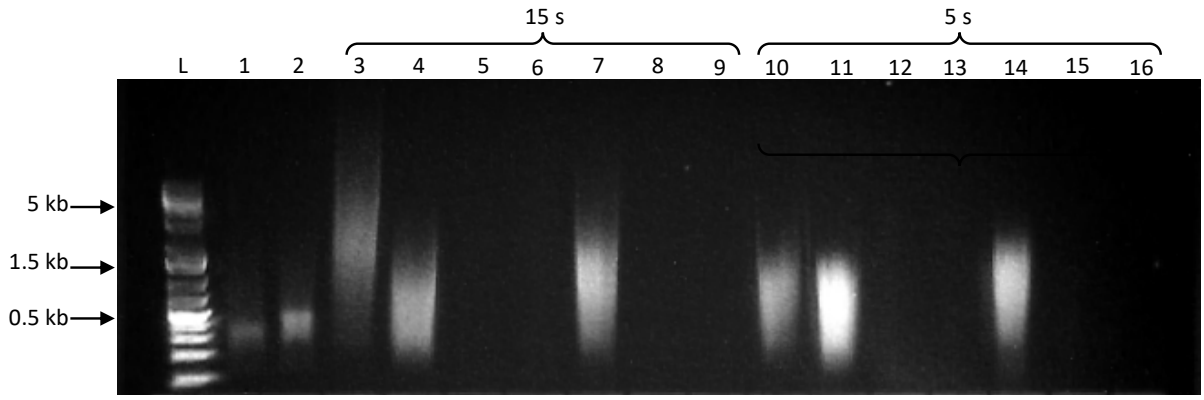


Figure 2.36. *Lonza™ agarose gel electrophoresis of qPCR products including negative controls (no primer, no duplex and no polymerase). Lanes 3-9 represent 15 second qPCR anneal time and lanes 10-16, 5 second qPCR anneal time. Lanes 1-2: Size recovered fractions before qPCR of fractions 3 and 4, respectively, Lanes 3-4: Amplified dA-GC fraction 3 and 4, Lanes 5-6: No primer control of fraction 3 and 4, Lane 7: No duplex control, Lanes 8-9: No polymerase control of fraction 3 and 4, Lanes 10-11: Amplified dA-GC fraction 3 and 4, Lanes 12-13: No primer control of fraction 3 and 4, Lane 14: No duplex control, Lanes 15-16: No polymerase control of fraction 3 and 4. L = DNA ladder.*

Lanes 1 and 2 represented the size recovered samples before qPCR and indicated the length at which the target DNA should be after qPCR, since it should have only amplified and not grown in length, lane 1 = 300 bp and lane 2 = 500 bp. For the 15 second anneal time, lanes 3 and 4 showed that after amplification the dA-GC samples had grown in length and created a large distribution of base pair lengths, fraction 3 extended to much longer lengths than fraction 4. It was also evident from these gel lanes, particularly lane 4, that the distribution of DNA started from lengths below that of the starting sample lengths. Lane 7 which contained the no duplex control showed a similar distribution of lengths to the DNA in lanes 3 and 4 and hence, explained the likeness of the qPCR traces and concentrations between the dA-GC extDNA samples and the working negative control.

The dA-GC qPCR products not only contained the target DNA at lengths of 300 and 400 bp, it also potentially contained self-annealed structures and large concatemeric forms of the target DNA due to the self-complementarity of the sequence. Hence, this could have created very long length DNA. Furthermore, since their qPCR traces and concentration values were so alike, it was reasonable to assume that the growth of the primer was also affecting the qPCR results of the target DNA. This would have contaminated the desired product and falsely elevated the absorbance values and qPCR fluorescence read outs.

The 5 second anneal qPCR products, lanes 10 - 16, showed the same results as those for the 15 second anneal and confirmed there was an issue with the primer. From these

findings, it was concluded that the dA-GC extDNA sample had amplified but contained a mixture of products; dA-GC extDNA and duplex primer DNA. To understand how the single stranded poly(T) primer formed duplex structures during the qPCR, snake venom digestion of the samples was performed followed by HPLC**, Figure 2.37**.

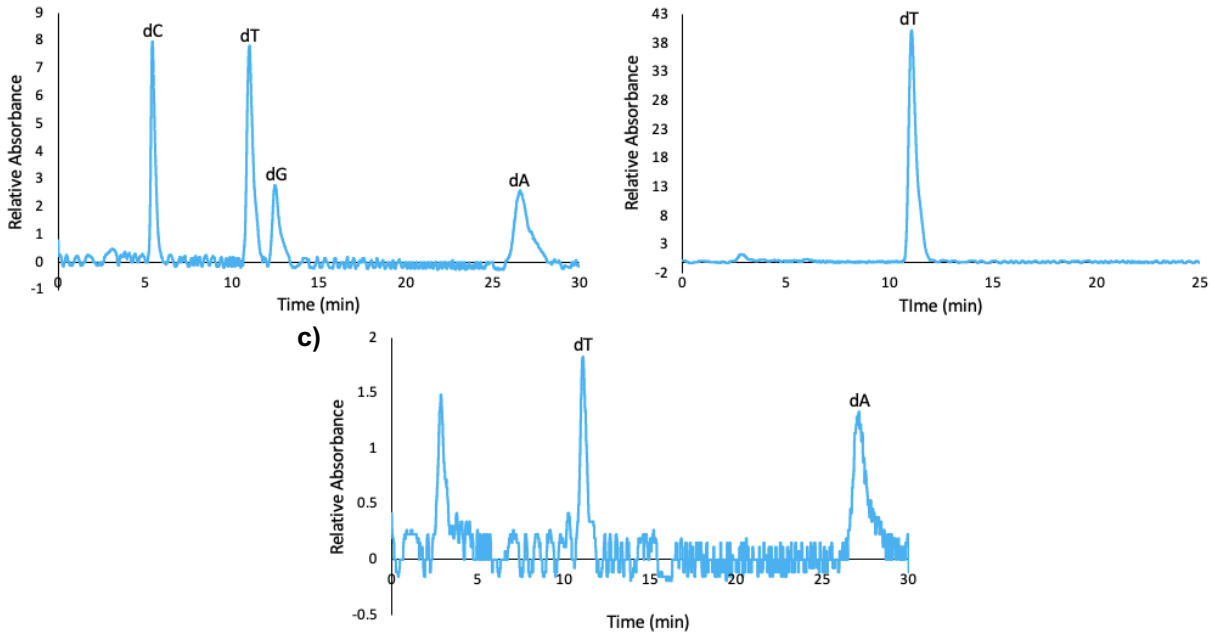


Figure 2.37. *HPLC analysis* ***a)*** *dNTP standards* ***b)*** *Primer T before qPCR* ***c)*** *Primer T after qPCR.*

dNTP standards were run in **Figure 2.37a** and indicated the retention times to detect on elution of the digested qPCR samples. The primer, poly(T), was injected, **Figure 2.37b**, and showed that it was pure, only dT bases present, and was not contaminated with any other bases before being used in the qPCR samples. Finally, the no duplex negative control sample after 40 qPCR heat-cool cycles was injected and it was evident that the product contained dT and dA bases from the presence of the new dA peak in the HPLC chromatogram. From the HPLC trace it was clear that the poly(T) primer had formed a duplex structure which allowed the polymerase to read the sequence and insert bases. It was determined that this occurred through the formation of T-T base mismatch, which allowed the primer to anneal with itself to form dsDNA and polymerase to insert dA bases into the gaps in the sequence. With the introduction of dA bases, slippage or Whitfield extension could have occurred and caused elongation of the DNA length, as observed in lanes 7 and 14 of the agarose gel, **Figure 2.36**.

Table 2.11. *HPLC retention times for qPCR primer T before and after qPCR.*

Sample	Retention Time (mins)			
	dC	dG	dT	dA
dNTP standards	5.5	11.5	13.4	29.1
Primer poly T – before	-	-	11.1	-
Primer poly T - after	-	-	11.1	27.2

The retention times, **Table 2.11**, for the HPLC samples showed an earlier elution time for bases dT and dA in the poly(T) samples before and after qPCR, compared to the dNTP standards. The dT was eluting at 11 minutes and dA at 27 minutes because there was no dG in the sample. This was because in the correct conditions, dG and dT are separate peaks on elution from the column and have different retention times. The elution of dG before dT and dA delays their elution and increases their retention time. However, in the incorrect conditions, dG and dT are eluted as one peak in the HPLC chromatogram, so it was not surprising that dT eluted earlier than in the dNTP standards.

After investigation of the primer behaviour during qPCR, it was realised that having dA present in the reaction mixture allowed for primer extension after self-annealing, since the polymerase could insert the dA against the dT when reading the duplex overhang. With the target DNA containing a poly(A) tail, there was no requirement to include dATP in the reaction mixture, since it did not need to be inserted by the polymerase to complete the target sequence. This would hopefully stop the primer from giving a positive response as it would be no longer possible for it to extend, even if T-T base mismatches occurred.

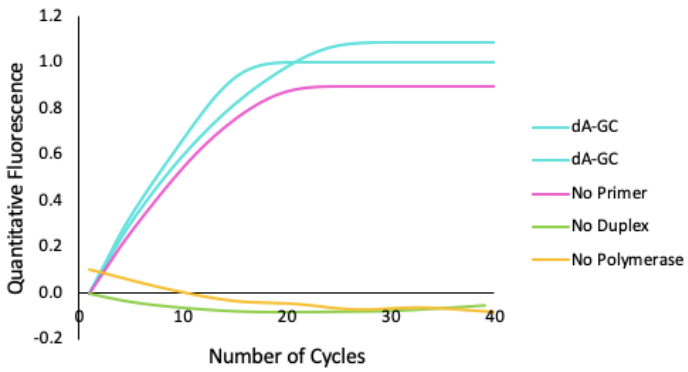


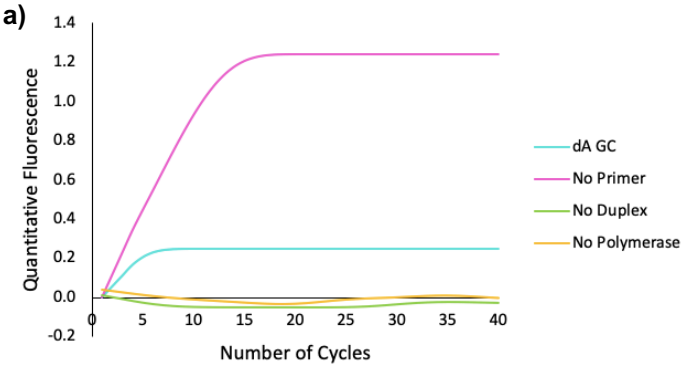
Figure 2.38. *qPCR using dA-tailed $[GC]_n/[CG]_n$ (not size recovered) with no dATP in the dNTP mixture.*

Initially, qPCR with no dATP base in the reaction mixture was attempted on a non-size recovered dA-$[GC]_n/[CG]_n$ sample, **Figure 2.38**. The qPCR trace showed that the dA-GC

sample increased in fluorescence indicating amplification of the target sample. However, the no primer negative control, sample which contained dA-GC DNA only, appeared to also increase in fluorescence with a trace that was similar to the target DNA. On the other hand, there was success in stopping the no duplex negative control from working and without the dATP base, the sample did not amplify and behaved as expected from a negative control.

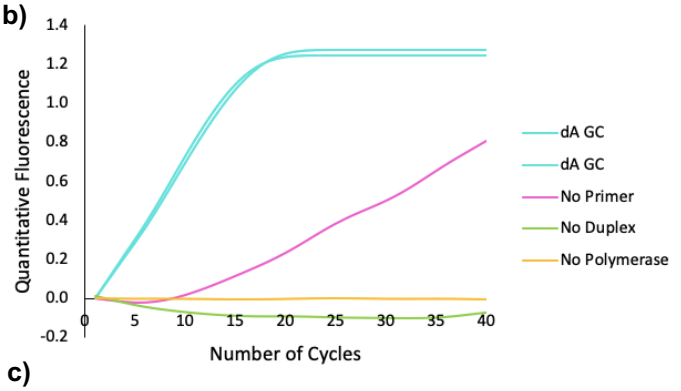
As previously discussed, salt concentration is important for the optimisation of PCR conditions. To try and achieve a situation where none of the negative controls produced positive responses, a study of the variation in $MgSO_4$ salt concentration on the qPCR products was undertaken. Only the concentration of $MgSO_4$ was changed to 2, 3 and 5 mM, all other qPCR conditions were kept constant.

a)



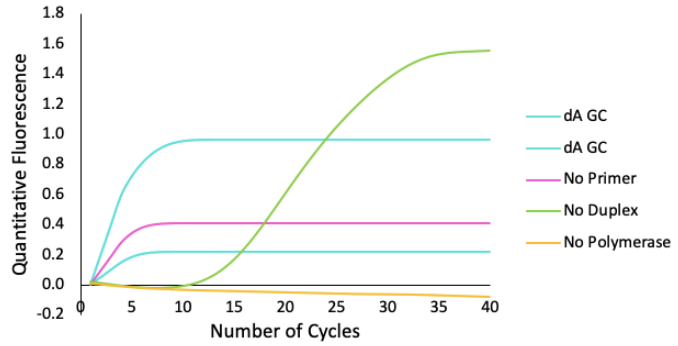
Sample	Abs @ 260 nm	Concentration (ng/µL)
dA-GC (F2)	0.5	25.0
No Primer	0.27	13.7
No Duplex	Salt	Salt
No Polymerase	Salt	Salt

b)



Sample	Abs @ 260 nm	Concentration (ng/µL)
dA-GC (F3)	0.22	10.9
dA-GC (F3)	0.30	15.2
No Primer	Salt	Salt
No Duplex	Salt	Salt
No Polymerase	Salt	Salt

c)



Sample	Abs @ 260 nm	Concentration (ng/µL)
dA-GC (F4)	0.44	22.1
dA-GC (F4)	0.59	29.5
No Primer	0.96	48.0
No Duplex	0.25	12.3
No Polymerase	Salt	Salt

Figure 2.39. *qPCR using size recovered dA-tailed $[GC]_n/[CG]_n$ with no dATP in the dNTP mixture and varying $MgSO_4$ concentrations.* ***a)*** *Fraction 2 (400 bp) and 2mM $MgSO_4$* ***b)*** *Fraction 3 (500 bp) and 3mM $MgSO_4$* ***c)*** *Fraction 4 (750 bp) and 5mM $MgSO_4$.*

From all three salt conditions, **Figure 2.39**, it was evident that the concentration of salt changed the outcome of the qPCR experiment, especially for the negative controls. **Figure 2.39a** used only one dA-GC target sample as there was not enough size recovered DNA at 400 bp to make more target samples as well as the negative controls. 2 mM $MgSO_4$ promoted the slippage of the no primer negative control as the fluorescence trace was 5x greater than the dA-GC target sample. Although, in terms of concentration a greater amount of target dA-GC DNA was produced, 25.0 compared to 13.7 ng/µL. This suggested the target DNA was a combination of no primer negative control and target DNA.

The second experiment used 3 mM $MgSO_4$, **Figure 2.39b**, and produced a more favourable outcome for the negative controls. As always, the no polymerase negative control did not

produce a response, and the no duplex negative control also did not increase in fluorescence under these conditions. The no primer negative control grew in a more linear fashion than it had done previously, where growth only started to occur around 10 cycles. Furthermore, the absorbance of the no primer negative control was not detected by UV-Vis spectroscopy and was recorded as salt due to the shape of the UV-Vis trace and the lack of 260 nm peak. Overall, the 3 mM $MgSO_4$ concentration produced a more positive result for the amplification of $[GC]_n/[CG]_n$ DNA.

The final $MgSO_4$ concentration studied was 5 mM and promoted the growth of two out of three of the negative controls; no primer and no duplex. These findings were also reflected in the concentration of the samples after qPCR and purification, as the samples which gave positive fluorescence traces also contained a 260 nm peak. Therefore, these conditions were disregarded as possible optimised conditions.

To visualise the findings from the investigation into the variation of $MgSO_4$ salt concentration on the qPCR products, a Lonza™ agarose gel of all three sets of salt conditions was performed, **Figure 2.40**.

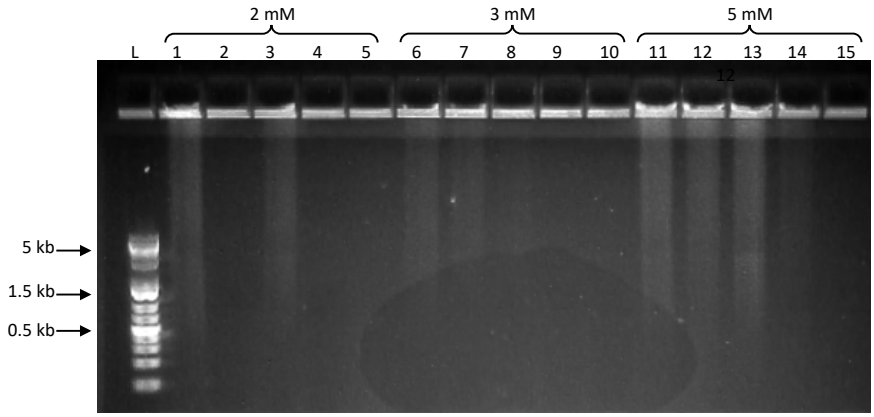


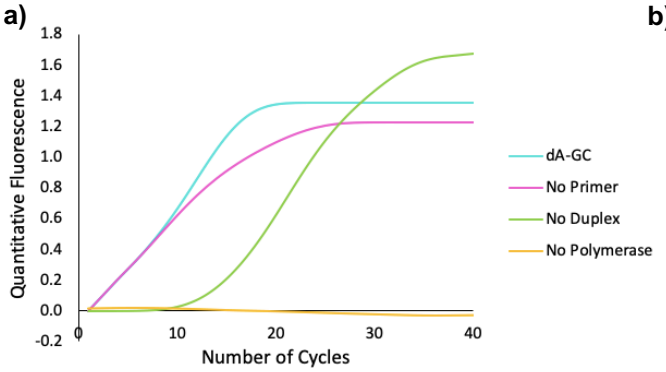
Figure 2.40. *Lonza™ agarose gel of qPCR samples with varying $MgSO_4$ concentrations and negative controls. Lanes 1-2: dA-tailed $[GC]_n/[CG]_n$ (200 bp) with 2 mM $MgSO_4$, Lane 3: No primer, Lane 4: No duplex, Lane 5: No polymerase, Lanes 6 - 7: dA-tailed $[GC]_n/[CG]_n$ (500 bp) with 3 mM $MgSO_4$, Lane 8: No primer, Lane 9: No duplex, Lane 10: No polymerase, Lanes 11 - 12: dA-tailed $[GC]_n/[CG]_n$ (750 bp) with 5 mM $MgSO_4$, Lane 13: No primer, Lane 14: No duplex and Lane 15: No polymerase. L = DNA ladder.*

The agarose gel showed that the purified DNA corresponded to the qPCR traces in **Figure 2.39**, where Lanes 1 to 5 were the 2 mM $MgSO_4$ samples, lanes 6 - 10 were the 3 mM $MgSO_4$ and lanes 11 - 15 were the 5 mM samples. All samples were trapped in the

wells of the gel which suggested the DNA had either grown to very long lengths or formed large concatemeric/self-annealed structures, meaning the DNA was unable to run through the pores of the gel as the structures were too large. Therefore, although some promise was shown with the 3 mM $MgSO_4$ concentration from the qPCR traces, the DNA did not retain its original size recovered length.

After taking into consideration the salt concentration of the reaction mixture, the next logical condition to investigate was the annealing temperature used during the qPCR. The annealing temperature is dependent on the Tm of the two hybridising single strands, in this case poly(T) to the poly(A) tail on the target DNA. As the length of the poly(A) tail on the size recovered DNA strand was unknown, the base pair length of the poly(T) primer strand was used to calculate the Tm value. The theoretical Tm was determined as 31.3 °C, based on the 20 bp oligo hybridisation using the theoretical OligoCalc programme. Since the theoretical Tm was below that of the annealing temperature used in the qPCR heat-cool cycles, 60 °C, the annealing of the primer DNA would not be favourable since the duplex would melt at this temperature rather than form. Therefore, it was decided that an annealing temperature of 26 °C would be used in order to increase the favourability of amplification of the target sample and decrease the chances of the negative controls producing a positive response, **Figure 2.41**.

a)



b)

Sample	Abs @ 260 nm	Concentration (ng/µL)
dA-GC (F4)	0.35	17.5
No Primer	0.27	13.3
No Duplex	Salt	Salt
No Polymerase	Salt	Salt

Figure 2.41. *qPCR using size recovered $[GC]_n/[CG]_n$ fraction 5 (1,000 bp) with no dATP in the dNTP mixture and 3 mM $MgSO_4$. qPCR anneal temperature reduced to 26 °C.*

It was clear that the reduction in the annealing temperature still promoted the amplification of the dA-GC target DNA. Furthermore, both the no primer and no duplex negative controls gave positive fluorescence signals; however, the no duplex negative control had a different fluorescence trace. This was due to the slower increase in fluorescence, indicating that duplex formation was low and there was little intercalation

of the Pico green dye until 20 cycles. These fluorescence results were reflected in the UV-Vis absorbance results, after DNA purification, as the dA-GC and the no primer negative control gave a small 260 nm peak, 0.35 and 0.27 respectively. The no duplex negative control had no detectable 260 nm peak and was recorded as salt, so the fluorescence trace was potentially produced from a very small concentration of DNA, which was below 100 bp. The DNA was therefore lost during PCR clean up and hence, gave no 260 nm peak in the UV-Vis spectrum.

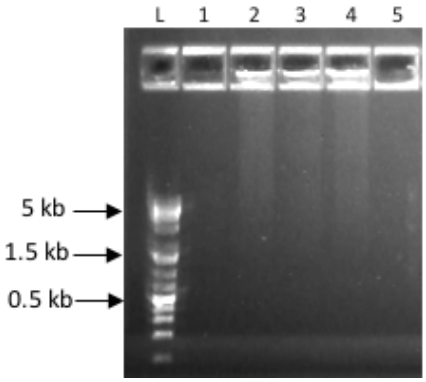


Figure 2.42. *Lonza™ agarose gel of qPCR samples with reduced annealing temperature (26 °C), no dATP in the dNTP mixture and 3mM $MgSO_4$. Lane 1: blank well lane 2: dA-tailed $[GC]_n/[CG]_n$ Lane 3: No primer, Lane 4: No duplex, Lane 5: No polymerase.*

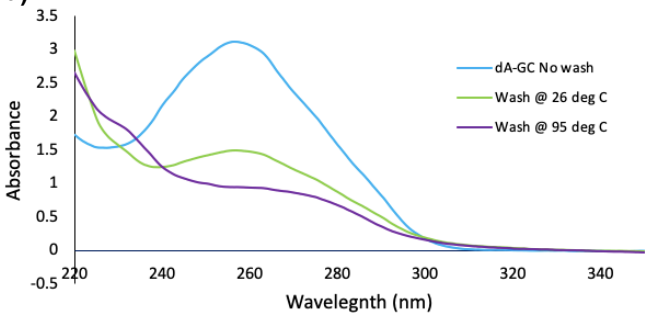
Moreover, an agarose gel was run of the qPCR samples with the lowered annealing temperature, **Figure 2.42**, and showed that DNA was present, not only in the dA-GC and no primer sample but also in the no duplex sample despite the lack of 260 nm peak in the UV-Vis trace. As in **Figure 2.40**, regardless if it was a target sample or a negative control, all DNA was trapped in the wells of the agarose gel due to the increase in length during qPCR which caused entanglement and networking of the DNA. Therefore, the DNA was no longer 1000 bp, as per the length after size recovery.

Even though the 3' end should have been blocked by the synthetically added dA-tail, it was theorised that since the DNA was elongating during the qPCR that some of the DNA was not being tailed. Therefore, the DNA could slip and extend during the qPCR as the reaction conditions and components were similar from those in the Whitfield enzymatic extension. Hence, the very long DNA observed in the agarose gels after qPCR in the dA-GC target samples and in the no primer negative control.

Although the dA-tailed DNA was purified with a PCR clean up kit after incubation, this did not remove any untailed DNA, only leftover reactants. A new purification method was

designed to remove any DNA which was not dA-tailed and would, hopefully, avoid any DNA extension from occurring during qPCR which skewed previous results. It was thought that a poly(T) 20mer on a CPG bead would be able to bind to the dA-tailed DNA strands and the unreacted DNA would remain in the solution. The poly(T) CPG beads were packed into a column and the dA-tailed DNA was passed through the column via a syringe. After being passed through the column at 26 °C, to facilitate annealing, the DNA was collected in another 1 mL syringe connected to other side of the column. This process was repeated three times and the wash solution collected, before the column was washed with water heated to 95 °C to remove the dA-tailed DNA bound to the CPG bead. The washes were concentrated to 20 μL for comparison to the concentration before purification by UV-Vis spectroscopy, **Figure 2.43**.

a)



b)

Sample	Abs @ 260 nm	Concentration (ng/μL)
dA-GC No Wash	3.07	153.5
Wash @ 26 °C	1.45	73.5
Wash @ 95 °C	Salt	Salt

Figure 2.43. *UV-vis spectroscopy from CPG bead poly(dT) purification with dA-tailed $[GC]_n/[CG]_n$.*

Analysis by UV-Vis spectroscopy showed that most of the DNA was removed in the 26 °C wash as the concentration was now 73.5 ng/μL compared to the starting concentration of 153.5 ng/μL. It was evident from the 95 °C wash that no DNA was recovered as the trace had no 260 nm peak and therefore, showed no sensible absorbance data for DNA, so was identified as salt. There was 80 ng/μL of DNA unaccounted for by UV-Vis spectroscopy and so qPCR, **Figure 2.44**, of the two washes was performed in case there was a small amount of DNA in the 95 °C wash hidden by the salt peak.

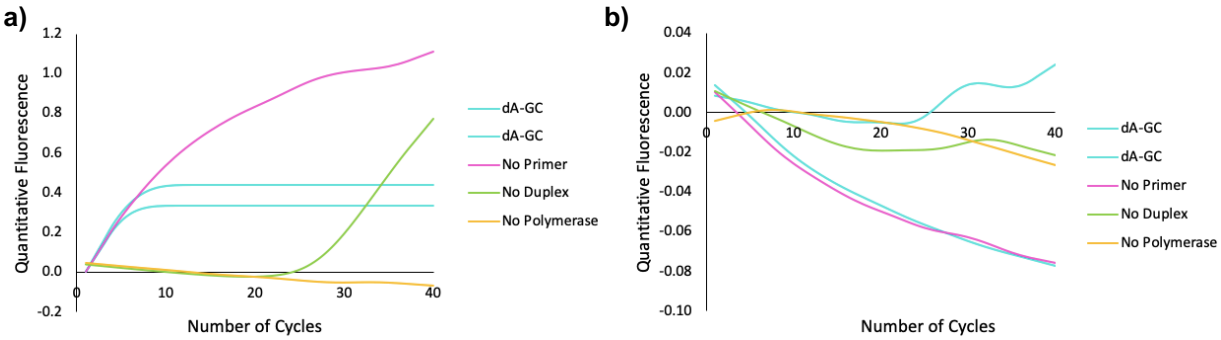


Figure 2.44. *qPCR using CPG bead poly(dT) purified dA-tailed $[GC]_n/[CG]_n$ DNA.* ***a)*** *DNA wash at 26 °C* ***b)*** *DNA wash at 95 °C.*

The 26 °C wash after qPCR, **Figure 2.44a**, further confirmed that there was DNA present in this sample and that there was no significant difference in how the qPCR performed in comparison to previous experiments. This was expected since the sample should have contained the un-reacted DNA from the dA-tailing protocol and would have extended during qPCR, thus producing a positive fluorescence trace. **Figure 2.44b**, validated the findings from the 95 °C UV-Vis data as all samples failed the qPCR experiment, since no trace gave a positive reading, therefore no DNA was present. From these findings, it was concluded that the dA-tailed DNA had either not bound to the CPG bead during the annealing step or was still on the CPG bead as it had not been denatured during the 95 °C wash. This was potentially due to limited denaturing time at 95 °C since the water cooled quickly during transfer between syringe-column-syringe.

The CPG-bead purification experiment was repeated and two extra 95 °C washes were added, along with using a PCR clean up kit on the washes, as this ensured that any reactants from the dA-tailing incubation was removed. Firstly, the column was washed with 95 °C water which would remove any residual DNA from the previous purification attempt. This was followed by the new purification protocol, a 26 °C wash and two 95 °C washes, with fresh dA-GC extDNA. The UV-Vis spectrum was recorded before and after using the PCR clean up kit **Figure 2.45**.

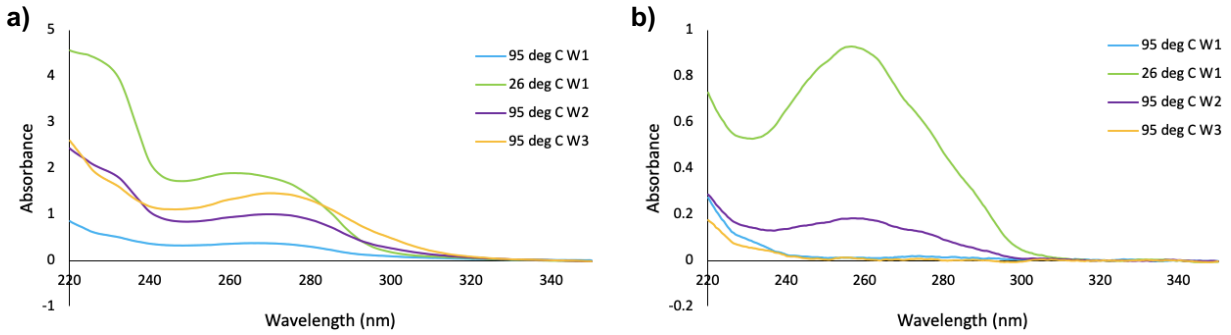


Figure 2.45. *UV-Vis spectroscopy a) CPG bead poly(dT) purified dA-tailed $[GC]_n/[CG]_n$ with both 26 °C and 95 °C washes. b) 2nd purification after CPG bead using QIAGEN column.*

From the UV-Vis traces before PCR purification, it was evident that there was not only DNA in the washes as the DNA peak was hidden by the residual dA-tailing reaction components peak. After PCR clean up, DNA peaks were evident in the 26 °C wash and the second 95 °C wash and so, these samples were taken forward to the qPCR, **Figure 2.46**.

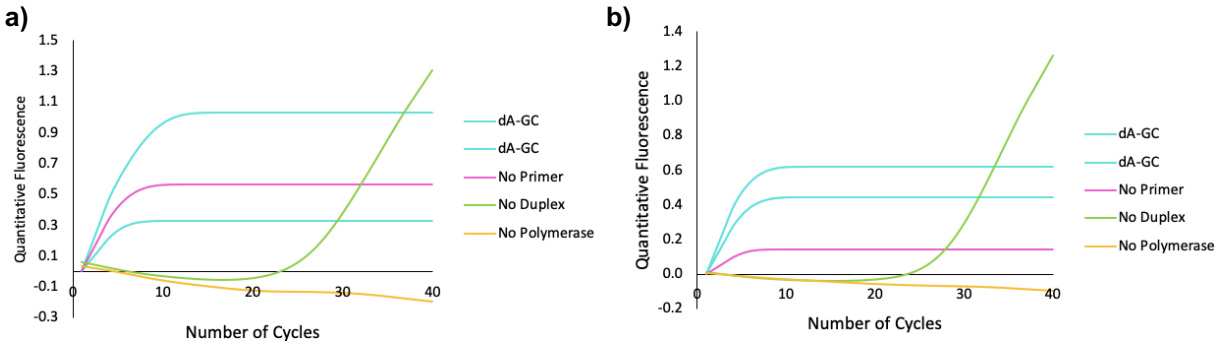


Figure 2.46. *qPCR after QIAGEN column purification.* ***a)*** *26 °C wash 1* ***b)*** *95 °C wash 2.*

After qPCR, it was clear that in both wash samples the no primer negative control continued to produce a positive fluorescence response. It was evident that the no primer negative control increase had reduced after two 95 °C washes which was potentially due to there being less un-reacted DNA, after the dA-tailing, being present in the sample. Furthermore, in both temperature samples, the no duplex negative control began to increase in fluorescence at approximately 25 cycles and although growth was slow, the fluorescence intensity was above that of the target DNA samples. However, the UV-Vis absorbance data showed that even though the fluorescence intensity was higher for the no duplex negative control, the concentration after purification was very low compared to the dA-GC target DNA products, **Table 2.12**.

Table 2.12 *Concentration comparison after qPCR and purification of dA-GC target DNA against the no duplex negative control for the 26 ℃ and 95 ℃ CPG washes.*

Wash Temperature	qPCR product Concentration (ng/μL)	
	dA-GC target DNA	No Duplex Negative Control
26 °C	233.27	37.71
	324.65	
95 °C	241.26	26.79
	245.84	

Overall, the CPG bead purification did not help to reduce the positive fluorescence responses of the negative controls, and therefore, offered no advantage for use in the amplification of size specific designer DNA over other methods discussed in this chapter.

2.3 Conclusions

In this chapter, the main aim was to produce aliquots of designer DNA which was of size-specific base pair length. These aliquots were produced but the concentration was not accurately amplified and required further investigation.

To improve the method of recovering small base pair length ranges of DNA, the speed at which the Lonza gel was run was slowed to 150 V. This allowed for slightly tighter bands of DNA to be recovered, although from the agarose gel electrophoresis, streaking of the DNA from one recovered fraction to its neighbouring fractions was still evident. The method was further improved by allowing the gel to equilibrate in order for the DNA to come to a complete stop before it was removed from the recovery wells, as it was noticed that after stopping the voltage the DNA continued to run for approximately 30 seconds. These recovered aliquots had small base pair length ranges of target DNA sequences and were confirmed to be of specific length by agarose gel electrophoresis, which showed tight bands of DNA post size recovery and therefore, an improvement from the original method.

Unfortunately, although the target sequence was size specific and no longer comprising of a large distribution, the concentration of these recovered DNA aliquots was low and could not be used for further experimental work, for example in DNA nanoassembly. In the interest of creating larger concentrations of size specific designer DNA, qPCR amplification was used alongside blunt end ligation and dA-tailing techniques to facilitate the replication of the target DNA sequence and length. The blunt end ligation technique used T4 DNA ligase and inserted a primer sequence onto the 3' end of the target DNA through the exposed 5' phosphate on the primer. The initial primer insert sequence proved to be of poor design as self-annealing sites and hairpins caused misleading fluorescent amplification traces in the qPCR. A new primer insert was designed which had no apparent self-annealing sites or potential hairpin formations, however problems with the replication of the target sequence were still present as the no duplex negative control, sample containing only primer DNA with other qPCR components, continued to give positive fluorescence traces during qPCR.

Moreover, to avoid issues with complicated primer design, the simple system of attaching a primer insert was employed by dA-tailing the 3' end of the target DNA. Characterisation of the successful dA-tailing was observed by agarose gel electrophoresis and HPLC

analysis. Initially, dATP was included in the qPCR reaction mixture, however it was found that it was causing the negative controls to produce concatemeric and self-annealed structures which could extend in length, hence producing positive qPCR responses and inaccurate amplification data. In order to counter act the negative controls giving a positive qPCR trace, it was realised that by limiting this method to sequences which contained only dG and dC, no dATP was required in the reaction mixture. By removing the dATP, the no duplex control behaved as expected for a negative control, however the no primer negative control increased in fluorescence values. Several strategies, such as reducing the annealing time and varying the $MgSO_4$ salt concentration, were investigated in order to reduce the effect of the no primer negative control on the replication and amplification of the target DNA. Furthermore, different purification methods were investigated by use of CPG bead and PCR clean up kit. The qPCR results showed no improvement, particularly in terms of the no primer negative control.

Although no method of qPCR with dA-tailing showed definitive successful amplification results, it was thought that condition set 5, **Table 2.13**, produced the best result for the potential amplification of the target DNA samples. Being able to fully achieve high concentrations of size specific designer DNA would be helpful to DNA users for future DNA nanotechnology applications as the cost of this DNA would be less than that which is commercially available, and no base pair length restrictions would be incurred.

Table 2.13. *Summary of qPCR condition sets used in this chapter and the outcome of the qPCR indicated by the concentration and negative control behaviour.*

Condition Set	DNA Sequence	Ligation Technique	qPCR Conditions	Concentration after qPCR and purification	Negative controls (✕ positive qPCR response, ✓ negative qPCR response)		
					No Primer	No Duplex	No Polymerase
1	$[AT]_{10}/[TA]_{10}$ (size recovered)	Blunt end ligation – primer sequence 1	95 °C for 120 s 95 °C for 15 s 60 °C for 15s 72 °C for 60 s	66.7 ng/μL	✕	✕	✓
2	$[AT]_{10}/[TA]_{10}$ (bulk)	Blunt end ligation – primer sequence 2	95 °C for 120 s 95 °C for 15 s 60 °C for 15s 72 °C for 60 s	154.9 ng/μL	✓	✕	✓
3	$[AT]_{10}/[TA]_{10}$ (size recovered)	dA-tailing PCR Clean UP	95 °C for 120 s 95 °C for 15 s 60 °C for 15s 72 °C for 60 s	122.2 ng/μL	✓	✕	✓
4	$[AT]_{10}/[TA]_{10}$ (size recovered)	dA-tailing PCR Clean UP	95 °C for 120 s 95 °C for 15 s 60 °C for 5 s 72 °C for 60 s	58.2 ng/μL	✓	✕	✓
5	$[GC]_{10}/[CG]_{10}$ (size recovered)	dA-tailing PCR Clean UP	95 °C for 120 s 95 °C for 15 s 60 °C for 5 s 72 °C for 60 s	66.0 ng/μL	✓	✕	✓
6	$[GC]_{10}/[CG]_{10}$ (size recovered)	dA-tailing PCR Clean UP	95 °C for 120 s 95 °C for 15 s 60 °C for 15 s 72 °C for 60 s	46.1 ng/μL	✓	✕	✓
7	$[GC]_{10}/[CG]_{10}$ (size recovered)	dA-tailing PCR Clean UP	95 °C for 120 s 95 °C for 15 s 60 °C for 15 s 72 °C for 60 s 2 mM $MgSO_4$ No dATP	25.0 ng/μL	✕	✓	✓
8	$[GC]_{10}/[CG]_{10}$ (size recovered)	dA-tailing PCR Clean UP	95 °C for 120 s 95 °C for 15 s 60 °C for 15 s 72 °C for 60 s 3 mM $MgSO_4$ No dATP	15.2 ng/μL	✕* * Linear growth after 5 cycles	✓	✓
9	$[GC]_{10}/[CG]_{10}$ (size recovered)	dA-tailing PCR Clean UP	95 °C for 120 s 95 °C for 15 s 60 °C for 15 s 72 °C for 60 s 5 mM $MgSO_4$ No dATP	29.5 ng/μL	✕	✕	✓
10	$[GC]_{10}/[CG]_{10}$ (size recovered)	dA-tailing PCR Clean UP	95 °C for 120 s 95 °C for 15 s 26 °C for 15 s 72 °C for 60 s No dATP	17.5 ng/μL	✕	✕* *Slow exponential trace, no 260 nm peak in UV-Vis	✓
11	$[GC]_{10}/[CG]_{10}$ (size recovered) 26 °C wash	dA-tailing CPG bead purification (method 2) PCR Clean UP	95 °C for 120 s 95 °C for 15 s 60 °C for 15 s 72 °C for 60 s No dATP	324.65 ng/μL	✕	✕* *Linear growth after 25 cycles	✓
12	$[GC]_{10}/[CG]_{10}$ (size recovered) 95 °C wash	dA-tailing CPG bead purification (method 2) PCR Clean UP	95 °C for 120 s 95 °C for 15 s 60 °C for 15 s 72 °C for 60 s No dATP	243.84 ng/μL	✕	✕* *Linear growth after 25 cycles	✓

2.4 Experimental

2.4.1 Materials

Deoxynucleotide Preparation

Deoxynucleotides, **Table 2.14**, were purchased from Eurofins (Ebersberg, Germany).

Table 2.14. *Purchased deoxynucleotides*

Name	Origin	Sequence
$[AT]_{10}$	Eurofins	5'-ATATATATATATATATATAT-3'
$[GC]_{10}$	Eurofins	5'-GCGCGCGCGCGCGCGCGCGC-3'
TTAA Primer	Eurofins	5'-Phosphate-TTAACGGCCAGATGGCCAGC-Spacer C3-3'
TTAA Primer Complementary	Eurofins	5'-GCTGGCCATCTGGCCGTTAA-3'
TTAA Primer (2)	Eurofins	5'-Phosphate-TTAACGGCCAGATCGTCAGC-Spacer C3-3'
TTAA Primer (2) Complementary	Eurofins	5'-GCTGACGATCTGGCCGTTAA-3'

Primer Template Annealing

The oligo seed duplexes, **Table 2.15**, were prepared by the addition of DNA annealing buffer (10 mM Hepes pH 7.5, 100 mM sodium chloride (NaCl) and 1 mM ethylenediaminetetraacetic acid (EDTA)) to 10 µL of each deoxynucleotide required. The sample was heated to 95 °C for 10 minutes and cooled slowly to room temperature (25 °C). The annealed duplexes were then stored at -20 °C.

Table 2.15. *Summary of annealed duplexes.*

Name	Sequence	Final Duplex Length (bp)
$[AT]_{10}$ and $[TA]_{10}$	$[AT]_{10}/[TA]_{10}$	20
$[GC]_{10}$ and $[CG]_{10}$	$[GC]_{10}/[CG]_{10}$	20
TTAA Primer/ TTAA Primer Complementary	5'-Phosphate- TTAACGGCCAGATGGCCAGC-Spacer C3-3'/5'-GCTGGCCATCTGGCCGTTAA-3'	20
TTAA Primer (2)/TTAA Primer (2) Complementary	5'-Phosphate-TTAACGGCCAGATCGTCAGC-Spacer C3-3'/5'-GCTGACGATCTGGCCGTTAA-3'	20

Expression and Purification of DNA Polymerases

Table 2.16. *Description of DNA Polymerases*

DNA Polymerase	Properties	Source
Thermococcus gorgonarius polymerase B (Pfu-Pol) Z3 exo-	Archaeal family B Polymerase Low fidelity variant with the 5'→3' exonuclease activity removed and alterations to the fingers domain	Gifted from Dr. Colette Whitfield (PDR)
Pyrococcus polymerase B Deep Vent exo-	Family B polymerase with 5'→3' exonuclease activity removed	New England BioLabs Inc.

2.4.2 Enzymatic Extension and Purification

0.2 µM DNA duplex, 200 nM polymerase, 10x reaction buffer (**Table 2.17**), 2mM magnesium sulfate ($MgSO_4$) and 0.5 mM deoxynucleotide triphosphates (dNTPs – dCTP, dATP, dTTP and dGTP) were mixed thoroughly and made up to the desired volume with nanopure water.

Thermocycling was performed using a VWR thermocycler by the method detailed below:

	30 seconds at 95 °C
N (number of cycles) x	30 seconds at 55 °C
	120 seconds at 72 °C

The products were cooled to 4 °C after the reaction and held at this temperature until removed from the thermocycler.

The extended DNA products were purified using either a QIAGEN QIAquick PCR purification kit (QIAGEN, Manchester) or a Monarch PCR and DNA clean up kit (NewEngland Bio Labs Inc.). Manufacturer's protocol was followed.

Table 2.17. *Polymerase Reaction Buffer*

Buffer	Components
10x reaction buffer	200 mM Tris-HCl (pH 8.8 at 25 °C), 100 mM ammonium sulfate ($(NH_4)_2SO_4$), 100 mM potassium chloride (KCl), 1 % Triton X-100, 1mg/mL bovine serum albumin (BSA) and 20 mM magnesium sulfate ($MgSO_4$)
1x ThermoPol	20 mM Tris-HCl (Ph 8.8 at 25 °C), 10 mM ammonium sulfate ($(NH_4)_2SO_4$), 10 mM potassium chloride (KCl), 2mM magnesium sulfate ($MgSO_4$) and 0.1 % Triton X-100

2.4.3 Agarose Gel Electrophoresis

Standard

DNA extension products were analysed by gel electrophoresis in TBE (tris, boric acid and EDTA) buffer. 1 % agarose (Melford Biolaboratories Ltd., Ipswich, UK) (1 g) was added to 1x TBE buffer (100 mL) and heated to dissolve. DNA samples were supplemented with 6x loading dye (2.5 % Ficoll-400, 11 mM EDTA, 3.3 mM Tris-HCl (pH 8 at 25 °C), 0.017 % sodium dodecyl sulfate (SDS) and 0.015 % bromophenol blue) (Thermo Scientific, UK) and GeneRuler 1 kilo base plus DNA ladder was used (Thermo Scientific, UK). The gels were run at 100 V, 100 mA, 10 W, for approximately 1 hour and post stained with 1 μg/μL

ethidium bromide solution for 30 minutes. The gel was visualised using an ultra-violet transilluminator and imaged using UVIproMW software, version 11.02.

Denaturing

5 μg of extDNA was placed on ice while the gel plate was set up in the fumehood. 0.27 g agarose (Melford Biolaboratories Ltd., Ipswich, UK) was added to 54 mL nanopore water and heated to dissolve. 7 mL 10x MOPS buffer and 11 mL 37 % formaldehyde solution (Sigma Aldrich, UK) was added to the gel solution and transferred to the gel tray to set. 10 μL extDNA sample and 10 uL sample buffer (20 mM MOPS, 1 mM EDTA, 5 mM NaAcetate, 50 % (v/v) formamide and 37 % (v/v) formaldehyde) were mixed on ice, heated to 60 °C for 5 minutes, cooled on ice for 5 minutes and 2 μL loading dye (50 mM NaOH, 2mM EDTA, 20 % Ficoll-400 and 0.06 % bromophenol blue) added. The gel was run in 1X MOPS buffer at 100 V, 100 mA, 10 W, for approximately 1 hour and used GeneRuler 1 kilo base plus DNA ladder (Thermo Scientific, UK) as the standard. The gel was washed (x2) with nanopore water in the jitterbug to remove formaldehyde and post stained with 10 % 10 mg/mL ethidium bromide in 1x MOPS buffer for 10 minutes. The gel was then visualised using an ultra-violet transilluminator and imaged using UVIproMW software, version 11.02.

Lonza™ Flashgel Electrophoresis

Lonza™ gel electrophoresis was performed using 1.2 % agarose pre-cast gel kit and run as stated in the manufacturer's handbook. DNA was recovered from the recovery wells by addition of QIAGEN elution buffer (QIAGEN, Manchester, UK) to the wells. The recovered DNA fractions were purified using a QIAquick PCR purification kit (250) following the manufacturer's protocol.

2.4.4 Ultra-violet – Visible (UV-Vis) Spectroscopy

UV-Vis spectroscopy was carried out using a Thermo Scientific NanodropONE. 1.5 μL of sample was placed onto the stage and analysed between 1100 and 190 nm.

2.4.5 Atomic Force Microscopy (AFM)

The top layer of the mica surface was cleaned using parcel tape. 2.5 μL of DNA sample was added to the 2.5 μL of 2mM magnesium chloride ($MgCl_2$) solution. The 5 μL mixture was

spread on to the mica surface using a micropipette. After 5 minutes, the surface was washed with 5 µL of nanopure water and nitrogen gas was passed over the surface to remove excess water. The sample was dried for 30 minutes in the lamina flow fumehood. Light microscopy was used to locate the sample and cantilever position. AFM images were collected using a Dimension V with a nanoscope controller (Veeco Instruments Inc., Metrology Group, Santa Barbara, CA). The tapping mode was used with an etched silicon tip (Tap 300 Al-G, 200 kHz, 40 N/m) on an isolation table. Nanoscope 7.00b19 software was used to acquire data.

2.4.6 Ligation

A T4 DNA ligase, rapid DNA ligation kit was used as per manufacturer's protocol (Thermo Fisher Scientific, UK).

2.4.7 qPCR

10 ng DNA duplex (vector DNA), 200 nM polymerase, 10x reaction buffer (**Table 2.17**), 5mM magnesium sulfate ($MgSO_4$), 500 nM DNA primer and 0.5 mM dNTPs (dCTP, dATP, dTTP and dGTP), 200x dilution Quanti-iT Pico Green (Invitrogen, Thermo Fisher, UK) were mixed thoroughly and made up to the desired volume with nanopure water.

Thermocycling was performed using a Corbett Rotor-Gene 6000 thermocycler by the method detailed below:

Step 1)	120 seconds at 95 °C
	15 seconds at 95 °C
Step 2)	
	15 seconds at 60 °C
40 cycles x	
	60 seconds at 72 °C

The products were cooled to 4 °C after the reaction.

The extended DNA products were purified using either a QIAGEN QIAquick PCR purification kit (QIAGEN, Manchester) or a Monarch PCR and DNA clean up kit (NewEngland Bio Labs Inc.). Manufacturer's protocol was followed.

2.4.8 dA-Tailing

NEBNext dA-tailing module was used as per manufacturer's protocol (New England BioLabs Inc.). dA-tailed samples were purified using either a QIAGEN QIAquick PCR purification kit (QIAGEN, Manchester) or a Monarch PCR and DNA clean up kit (NewEngland Bio Labs Inc.). Manufacturer's protocol was followed.

2.4.9 DNA Digestion

0.2 mg snake venom phosphodiesterase, 100 units bacterial alkaline phosphate (BAP) in 10 mM magnesium chloride (pH 7.0) and 10 mM magnesium chloride were added to the dsDNA sample (1-2 absorbance units at 260 nm) and made to a volume of 100 μL with nanopore water. The sample was incubated at 37 °C for 16-18 hours. 50 μL of digested product was analysed by reverse phase HPLC as per **section 2.4.10**.

2.4.10 High Performance Liquid Chromatography (HPLC)

HPLC was performed on a Waters 2487 machine using Picolog software, and a Varian Pro Star using a C18 reverse phase column for both systems. Buffer C was 0.1 M triethyl ammonium acetate at pH 6.5 with 2.5 % acetonitrile, and buffer B was 0.1 M triethyl ammonium acetate at pH 6.5 with 65 % acetonitrile at the gradient 0-30 % B over a period of 25 minutes, increasing 10 % buffer B at 5-minutes. The gradient was returned to 0% B after 25 minutes at a rate of 1ml/min. Absorbance was monitored at a wavelength of 254 nm.

2.4.11 CPG Bead Purification

Poly(T) CPG beads were inserted into a column and the dA-tailed DNA solution was added to 0.5 mL of nanopore water heated to 26 °C. This solution was passed through the column into another syringe to facilitate annealing of dA-tailed DNA to the CPG bead. This was carried out 3 times back and forth between the two syringes holding the CPG bead column. The 26 °C wash was collected before washing the column three times with water heated to 95 °C. The washes were concentrated to 20 μL using a SpeedVac Concentrator. Additional 95 °C washes were introduced and were performed in the same manner as described above.

2.5 References

1. Whitfield, C. J., Turley, A. T., Tuite, E. M., Connolly, B. A. & Pike, A. R. Enzymatic Method for the Synthesis of Long DNA Sequences with Multiple Repeat Units. *Angew. Chem. Int. Ed.* **54**, 8971–8974 (2015).
2. Pettersson, U., Mulder, C., Deluis, H. & Sharp, P. A. Cleavage of adenovirus type 2 DNA into six unique fragments by endonuclease R-RI. *Proc. Natl. Acad. Sci. U. S. A.* **70**, 200-204 (1973).
3. Thuring, R. W. J., Sanders, J. P. M. & Borst, P. A freeze-squeeze method for recovering long DNA from agarose gels. *Anal. Biochem.* **66**, 213–220 (1975).
4. Blin, N., von Gabain, A. & Bujard, H. Isolation of large molecular weight DNA from agarose gels for further digestion by restriction enzymes. *FEBS Lett.* **53**, 84–86 (1975).
5. Fuke, M. & Thomas, C. A. Isolation of open-circular DNA molecules by retention in agar gels. *J. Mol. Biol.* **52**, 395–397 (1970).
6. Yang, R. C.-A., Lis, J. & Wu, R. Elution of DNA from agarose gels after electrophoresis. *Methods Enzymol.* **68**, 176–182 (1979).
7. Tabak, H. F. & Flavell, R. A. A Method for the Recovery of DNA from Agarose Gels. *Nucleic Acids Res.* **5**, 2321–2331 (1978).
8. Dretzen, G., Bellard, M., Sassone-Corsi, P. & Chambon, P. A reliable method for the recovery of DNA fragments from agarose and acrylamide gels. *Anal. Biochem.* **112**, 295-298 (1981).
9. Hengen, P. N. Methods and reagents: Recovering DNA from agarose gels. *Trends Biochem. Sci.* **19**, 388–389 (1994).
10. Smith, H. O. *[46] Recovery of DNA from gels*. *Methods in enzymology.* **65**, (Academic Press, 1980).
11. Downey, N. Extraction of DNA from Agarose Gels. in *E. coli Plasmid Vectors* 137-140 (Humana Press, 2003).
12. Extraction of DNA fragments from polyacrylamide gels using the QIAquick Gel Extraction Kit - QIAGEN. Available at: https://www.qiagen.com/gb/resources/resourcedetail?id=1426dbb4-da09-487c-ae01-c587c2be14c3&lang=en. (Accessed: 24th October 2018)
13. Schmitter, J., Mechulam, Y., Fayat, G. & Anselme, M. Rapid purification of DNA fragments by high-performance size-exclusion chromatography. *J. Chromatogr. B. Biomed. Sci. App.* **378**, 462–466 (1986).

14. Kotlyar, A. B. *et al.* Long, monomolecular guanine-based nanowires. *Adv. Mater.* **17**, 1901–1905 (2005).

15. Kotlyar, A. Synthesis of Long DNA-Based Nanowires. in *DNA Nanotechnology - Springer* **749**, 115–140 (2011).

16. Borovok, N., Molotsky, T., Ghabboun, J., Porath, D. & Kotlyar, A. Efficient procedure of preparation and properties of long uniform G4–DNA nanowires. *Anal. Biochem.* **374**, 71–78 (2008).

17. Little, R. C., Whitfield, C. J., Tuite, E. M. & Pike, A. R. The Synthesis of Designer DNA. in *DNA Nanotechnology: Methods and Protocols* (ed. Zuccheri, G.) 11–21 (Springer New York, 2018).

18. Jozwiakowski, S. K., Keith, B. J., Gilroy, L., Doherty, A. J. & Connolly, B. A. An archaeal family-B DNA polymerase variant able to replicate past DNA damage: occurrence of replicative and translesion synthesis polymerases within the B family. *Nucleic Acids Res.* **42**, 9949–63 (2014).

19. Schindelin, J., Rueden, C. T., Hiner, M. C. & Eliceiri, K. W. The ImageJ ecosystem: An open platform for biomedical image analysis. *Mol. Reprod. Dev.* **82**, 518–29 (2015).

20. Whitfield, C. J. Enzymatic Protocols for the Synthesis of Designer DNA. (Newcastle University, 2016).

21. Haugstad, G. Overview of AFM. *At. Force Microsc.* 1–32 (2012).

22. Jagtap, R. N. & Ambre, A. H. Overview literature on atomic force microscopy (AFM): Basics and its important applications for polymer characterization. *Indian J. Eng. Mater. Sci.* **13**, 368–384 (2006).

23. Pastré, D. *et al.* Adsorption of DNA to Mica Mediated by Divalent Counterions: A Theoretical and Experimental Study. *Biophys. J.* **85**, 2507–2518 (2003).

24. Nazari, Z. E. & Gurevich, L. Molecular Combing of DNA: Methods and Applications. *J. Self-Assem. Mol. Electron. SAME* **1**, 125–148 (2013).

25. Hansma, H. G., Revenko, I., Kim, K. & Laney, D. E. Atomic force microscopy of long and short double-stranded, single-stranded and triple-stranded nucleic acids. *Nucleic Acids Res.* **24**, 713–20 (1996).

26. Beaucage, S. L. & Iyer, R. P. Advances in the Synthesis of Oligonucleotides by the Phosphoramidite Approach. *Tetrahedron* **48**, 2223–2311 (1992).

27. Liu, Z. G. & Schwartz, L. M. An efficient method for blunt-end ligation of PCR products. *BioTechniques* **12**, 28, 30 (1992).

28. Liu, Z. G. & Schwartz, L. M. An efficient method for blunt-end ligation of PCR products. *BioTechniques* **12**, 28, 30 (1992).

29. Michelsen, B. K. Transformation of Escherichia coli Increases 260-Fold upon Inactivation of T4 DNA Ligase. *Anal. Biochem.* **225**, 172–174 (1995).

30. Real-Time vs Digital vs Traditional PCR - UK. Available at: 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. (Accessed: 11th February 2019)

31. Heid, C. A., Stevens, J., Livak, K. J. & Williams, P. M. Real time quantitative PCR. *Genome Res.* **6**, 986–94 (1996).

32. Gibson, U. E., Heid, C. A. & Williams, P. M. A novel method for real time quantitative RT-PCR. *Genome Res.* **6**, 995–1001 (1996).

33. Dragan, A. I. *et al.* Characterization of PicoGreen Interaction with dsDNA and the Origin of Its Fluorescence Enhancement upon Binding. *Biophys. J.* **99**, 3010–3019 (2010).

34. Dragan, A. I. *et al.* Metal-enhanced PicoGreen fluorescence: Application for double-stranded DNA quantification. *Anal. Biochem.* **396**, 8–12 (2010).

35. Wang, Y., Schellenberg, H., Walhorn, V., Toensing, K. & Anselmetti, D. Binding mechanism of PicoGreen to DNA characterized by magnetic tweezers and fluorescence spectroscopy. *Eur. Biophys. J.* **46**, 561–566 (2017).

36. Singer, V. L., Jones, L. J., Yue, S. T. & Haugland, R. P. Characterization of PicoGreen Reagent and Development of a Fluorescence-Based Solution Assay for Double-Stranded DNA Quantitation. *Anal. Biochem.* **249**, 228–238 (1997).

37. Dragan, A. I. *et al.* Metal-enhanced PicoGreen fluorescence: Application for double-stranded DNA quantification. *Anal. Biochem.* **396**, 8–12 (2010).

38. Panjkovich, A. & Melo, F. Comparison of different melting temperature calculation methods for short DNA sequences. *Bioinformatics* **21**, 711–722 (2005).

39. Preparation of Denaturing Agarose Gels | National Diagnostics. Available at: https://www.nationaldiagnostics.com/electrophoresis/article/preparation-denaturing-agarose-gels. (Accessed: 24th October 2018)

40. Kibbe, W. A. OligoCalc: an online oligonucleotide properties calculator. *Nucleic Acids Res.* **35**, 43–46 (2007).

41. Clark, J. M., Joyce, C. M. & Beardsley, G. P. Novel blunt-end addition reactions catalyzed by DNA polymerase I of Escherichia coli. *J. Mol. Biol.* **198**, 123–127 (1987).

42. NEBNext® dA-Tailing Module | NEB. Available at: https://www.neb.com/products/e6053-nebnext-da-tailing-module#Product Information. (Accessed: 24th October 2018)
43. Hashim, O. H. & Adnan, N. A. Coenzyme, cofactor and prosthetic group — Ambiguous biochemical jargon. *Biochem. Educ.* **22**, 93–94 (1994).

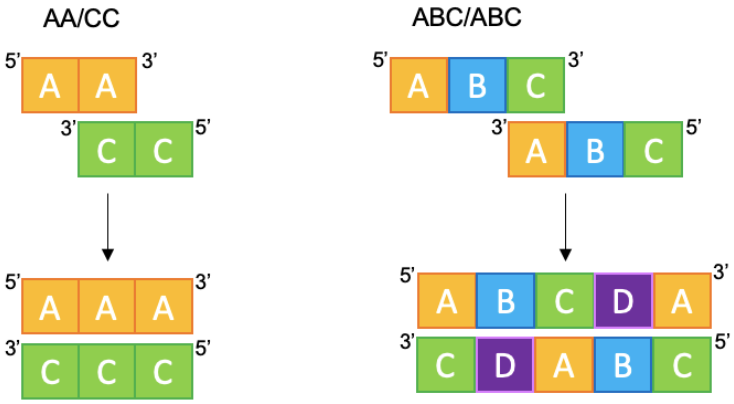
Chapter 3

ABCD-DNA: Synthesis of Designer DNA 2.0

Contents

3.1 Introduction

The Whitfield enzymatic extension method was used in **Chapter 2** to produce DNA that is not restricted in length, sequence or modification. This enzymatic method uses oligoseeds, which are made up of the same sequence that is repeated at least twice, *i.e.* a 3'-AA-5'/5'-CC-3' structure, where A is one sequence and C is its' complement. In order to expand the method, to fabricate DNA consisting of more than one sequence, 3'-ABC-5'/5'-ABC-3' oligoseed approach was proposed. A, B and C are different sequences of more than eight bases, where A and C are complementary. The proposed ABCD-DNA method introduces a new sequence D, complementary to B, into the DNA during enzymatic extension. This allows for an improvement on sequence spacing and affords more control over the deposition of materials at different sequence sites, **Scheme 3.1**.

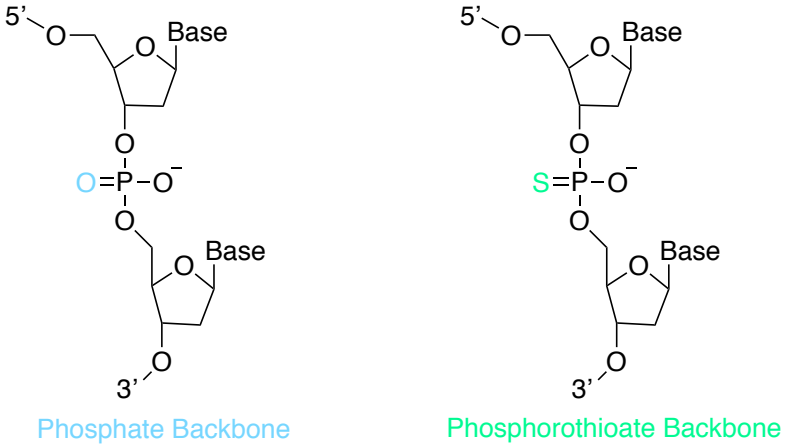


Scheme 3.1. *Initial AA/CC oligoseed structure used in the Whitfield enzymatic extension compared to the proposed new 3'-ABC-5'/5'-ABC-3' oligoseed structure.*

Chapter 3 explores a possible way of exploiting this DNA synthesis method by the incorporation of modifications within one or more of the sequences, A, B, C or D, for site specific binding. Previous research introduced in **Chapter 1**, detailed the many different base modifications which could be incorporated by the Whitfield enzymatic extension, such as single atom exchange modifications (5-I-dCTP and 5-Br-dUTP), modified purines (7-deaza-7-I-dATP and 6-S-dGTP) and long carbon chain introduction (5-octadiynyl-dCTP).[1]

During enzymatic synthesis using modified bases, the modifications are unique to each base and it is not possible to fabricate DNA of a mixed base sequence that bears the identical modification throughout the sequence. Therefore, binding sites which are not determined by the base must be through modifications of the DNA backbone, since is it identical between all bases in the sequence. The phosphorothioate functional group is a

good example of this modification, and is achieved by changing one of the non-bridging phosphate oxygens to a sulfur atom, **Scheme 3.2**.[2,3]



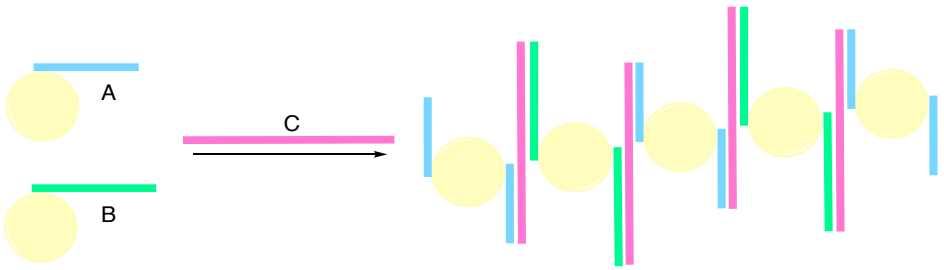
Scheme 3.2. *Phosphate backbone compared to phosphorothioate backbone modification.*

Generally, this modification has been used to produce antisense oligonucleotides for therapeutics due to their metabolic stability[4] and bioavailability.[3] More recently, phosphorothioates have been used to direct the assembly DNA to AuNPs (5 - 30 nm in diameter) to form spatially aligned particles or DNA-AuNP conjugates.[5–8] The popularity of incorporating the phosphorothioate modification within the DNA backbone is due to the fact that it maintains the structure of B-DNA, while offering a stronger binding affinity to soft metals.[9] Furthermore, it is advantageous over the terminal thiol-DNA to Au linkage, described in **Chapter 1**, as it is less expensive and can be located within the oligonucleotide sequence instead of at the 5' or 3' ends.[10]

Oligonucleotides containing the phosphorothioate modification can be synthesised by automated synthesis, although this approach is limited in length and number of modifications. Alternatively, Ghabboun *et al.*[11] showed that the phosphorothioate modification can be enzymatically incorporated to produce DNA of lengths longer than those commercially available by automated synthesis.[12] The phosphorothioate modified DNA produced showed no difference in height or appearance to that of standard DNA consisting of a phosphate backbone by AFM analysis.

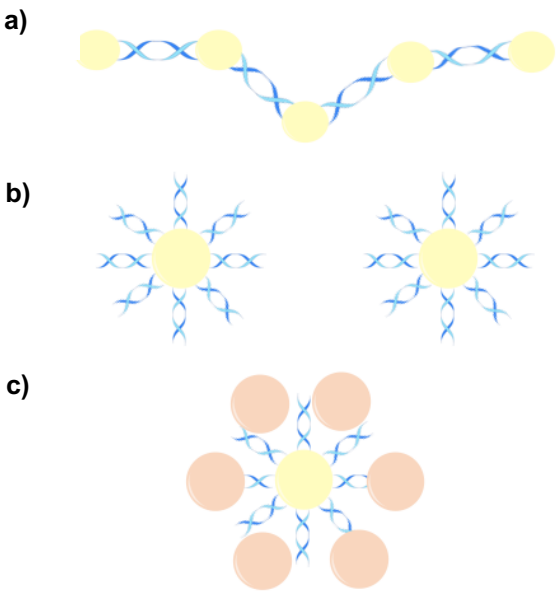
Phosphorothioate modified DNA oligos have been exploited for the directed assembley of AuNPs. Jiang *et al.*[10] loaded phosphorothioate oligos in parallel onto AuNPs by exploiting the DNA-AuNP interaction that is dominated by the sulfur atom.[13] 13 nm AuNPs were modified with short phosphorothioate oligos (A and B in **Scheme 3.3**) to generate two

sets of AuNPs. The A and B sets of AuNPs were then linked by hybridisation onto one ssDNA strand (C), where each short oligo was complementary to a section of the hybridising DNA strand. This meant that the AuNPs were spatially separated by strand C, **Scheme 3.3**.



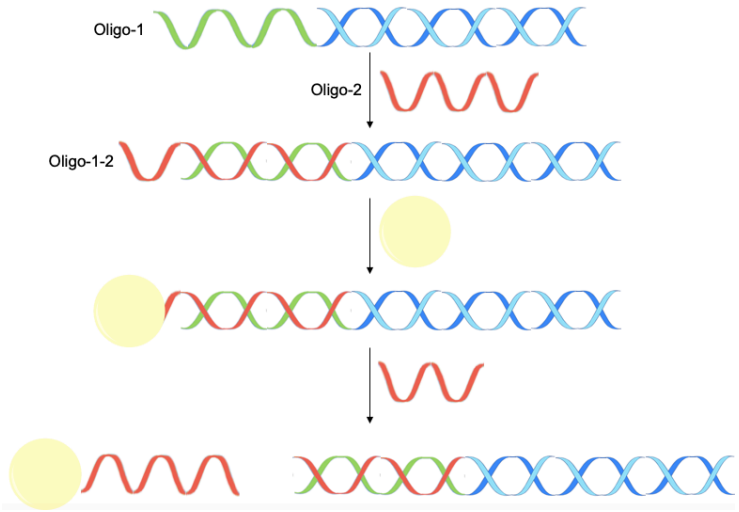
Scheme 3.3. *Phosphorothioate oligo modification of 13 nm AuNP surfaces to produce parallel AuNP-DNA structures by the hybridisation of various DNA oligo lengths.*[10]

Lubitz *et al.*[14] reported the use of phosphorothioate oligos to produce G-quadruplex (G4) DNA coated AuNPs. In order to attach the DNA to the AuNP, they used G4 DNA which contained flanking regions of phosphorothioated A bases on either end of the DNA. The modified G4 DNA was incubated with 15 nm AuNPs and at low DNA concentrations long nanoparticle chains were formed, however at high DNA concentrations fully coated AuNPs are produced. The coated AuNPs can then bind to citrate stabilised AuNPs and form flower like structures, **Scheme 3.4**.[14]



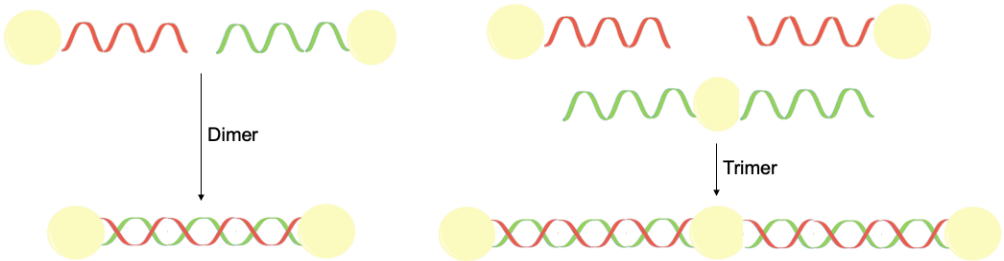
Scheme 3.4. a) *Formation of nanoparticle chains at low DNA concentrations* ***b)*** *Phosphorothiate G4-DNA coated AuNPs* ***c****) AuNP flower structure formed by the binding of citrate stabilised AuNPs to phosphorothioate G4-DNA coated AuNPs.*[14]

The linear nanoparticle chains formed at low DNA concentrations, **Scheme 3.4a**, produced spatially separated AuNPs with a distance equivalent to the length of the oligo used. Further AuNP spacing was achieved by the Kotlyar group by hybridising a double stranded oligo with a 25-base overhang (oligo-1) to a single stranded oligo containing an overhang of 10 phosphorothioate A bases (oligo-2).[15] The newly formed dsDNA (oligo-1-2) was incubated with AuNPs, to form a DNA-AuNP hybrid. Oligo-1 was separated from the complex by heat treatment in the presence of a short DNA complement, to generate the AuNP complexed with oligo-2, **Scheme 3.5**.



Scheme 3.5. *Synthesis of DNA-AuNP hybrids by multiple oligo hybridisations.*[15]

This approach can be used to make multiple variations of AuNP-oligo hybrid structures and allows for two particles containing complementary DNA strands to form a dsDNA bridge between the particles. Complexes containing two oligo strands on the surface will hybridise to another two particles to create a linear trimer, **Scheme 3.6**. These structures would again create spacing dependent on the length of the bridging DNA.

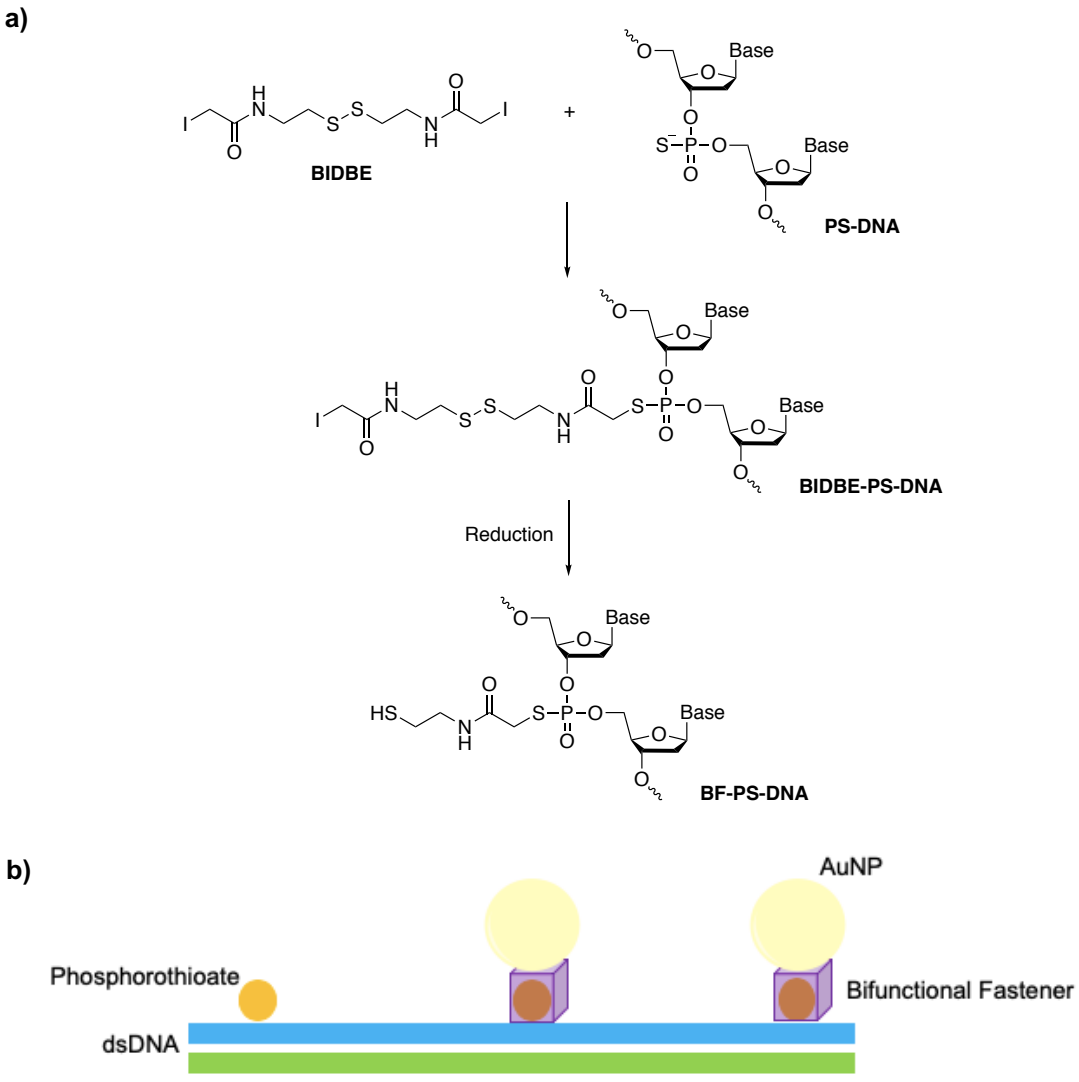


Scheme 3.6. *Dimer and trimer formation by complementary AuNP-oligo complex hybridisation.*[15]

There have not been many published examples of the assembly of AuNPs along a DNA strand using phosphorothioate modified DNA to control the position and the distance

between the nanoparticles. Lee *et al.*[16] reported the ability to achieve this by the addition of a bifunctional fastener at the phosphorothioate sites on the DNA oligo.

Automated synthesis was used to control the number and position of the phosphorothioate modifications programmed into the resulting oligo.[16] The bifunctional fastener contained an alkane thiol group which attached to the AuNPs and an iodoacetamine functionality on the other end which bound to the phosphorothioate group on the modified backbone. This enabled nanoparticle assembly at any modified backbone site where the position and distance between the nanoparticles was controlled without further DNA functionalisation of the AuNPs.



Scheme 3.7. a) *Mechanism for bifunctional fastener (BF) modification of phosphorothioate DNA (PS-DNA)* ***b)*** *Controlled assmebly of AuNPs on the modified DNA backbone.*[16]

The methods described above for the spatial binding of AuNPs to phosphorothioate modified DNA are focussed on short oligo sequences. There have also been some reports

of long, enzymatically produced phosphorothioated DNA binding to AuNPs.[17] However, it would be advantageous to develop a method to spatially place AuNPs in a linear fashion along a DNA guide through sequences of phosphorothioate linkages spaced in a controlled manner.

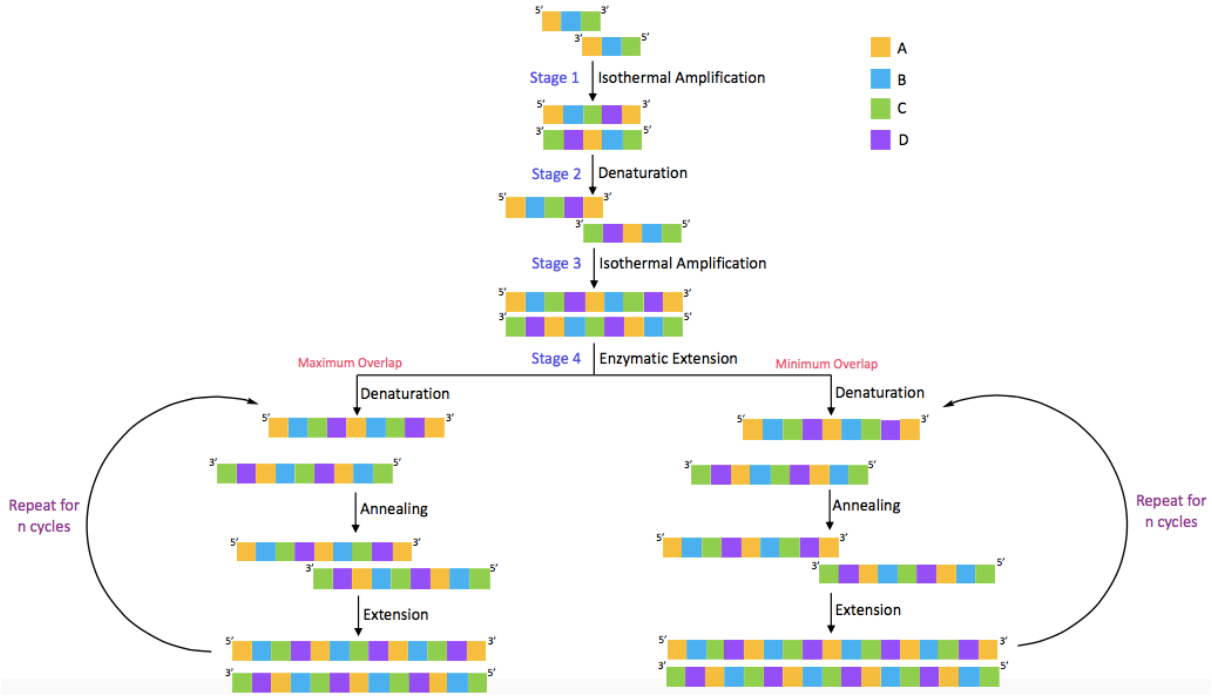
3.2 Results and Discussion

In this chapter, the versatility of the Whitfield enzymatic extension method is explored via the synthesis of long DNA products containing three multi-base sequences, ABC. This approach afforded the enzymatic incorporation of a fourth sequence, D and is defined as the ABCD-DNA method. Various novel designer DNA sequences, **Table 3.1**, were extended using the ABCD-DNA method to afford DNA lengths ranging from 500 - 5,000 bp. The method exploited the incorporation of the new sequence D, to allow for the successful inclusion of spatially separated sequences of phosphorothioate bases into ssDNA or dsDNA. This chapter details binding studies using 3 nm AuNPs to determine if the modifications successfully acted as spatial binding sites, affording AuNP separation.

Table 3.1. *ABCD-DNA oligoseed sequences.*

Name	Origin	Sequence
DNA-1	Eurofins	5'-AAAAAAAAAGCCCCCCCCCCTTTTTTTTTC-3' / 5'-CTTTTTTTTTCCCCCCCCCCGAAAAAAAAA-3'
DNA-2	Eurofins	5'-AAAATAAAATCCCCCCCCCCTTTTATTTTA-3' / 5'-ATTTTATTTTCCCCCCCCCCTAAAATAAAA-3'
DNA-3	Eurofins	5'-AAAATAAAATCCCCCCCCCCCCCCCCCCCCTTTTATTTTA-3' / 5'-ATTTTATTTTCCCCCCCCCCTAAAATAAAA-3'
DNA-4	Eurofins	5'-CCCCCCCCCCAAAAAAAAAAGGGGGGGGGG-3' / 5'-GGGGGGGGGGAAAAAAAAAACCCCCCCCCC-3'

The ABCD-DNA method used throughout this chapter is illustrated in Scheme **3.8**; this approach pairs an isothermal amplification step with the previously reported Whitfield enzymatic extension of user-defined oligoseeds. Here, the oligoseed consists of two partially complementary oligonucleotides, each divided into three sections; A, B and C, where the letters represent three separate sequences of 10 bases or more to ensure a stable overlap is achieved during annealing. The two oligonucleotides must be the inverse of one another; one sequence must read 5'-ABC-3', with the other sequence reading 5'-CBA-3' where section A is complementary to section C. When sections A and C hybridise, overhangs of AB or BC allow for the incorporation of a new set of bases complementary to section B, section D.



Scheme 3.8. *Schematic representation of ABCD-DNA method of DNA synthesis. Starting oligo must contain repeats of 10 bases, with the complementary strand the inverse to the initial oligonucleotide. The two mirroring oligonucleotides undergo two stages of isothermal amplification at 65 °C using Bst 2.0 polymerase to increase the stability of the overlap by creating full complementarity in the dsDNA. The enzymatic extension reaction mixture, including Z3 (exo-) polymerase, is added to the isothermal extension product and the dsDNA is extended using a user-defined number of cycles, n (where n=5, 10, 15 or 20).*

3.2.1 DNA-1 [$A_9GC_{10}T_9C$]

The Whitfield enzymatic extension process was employed as outlined in **Chapter 1**. The oligoseed [$A_9GC_{10}T_9C$]/[$A_9GC_{10}T_9C$] was called DNA-1, where A was A_9G, B was C_{10} and C was T_9C and was enzymatically extended with *Thermococcus gorgonarius* polymerase B (Pfu-Pol) Z3 exo$^-$ for 5, 10, 20 and 30 heat-cool cycles, to incorporate the new base set D, G_{10}. This determined the behaviour of the oligoseed during various heat-cool cycle lengths and whether this approach to synthesise long, spatially separated DNA was possible. Furthermore, if incorporation was successful, it also identified the optimum number of cycles for the production of long length extDNA with long repeats of difficult bases which are hard to incorporate via automatic synthesis. Initial characterisation of the DNA-1 extDNA product was performed by agarose gel electrophoresis and Image J analysis, **Figure 3.1**.

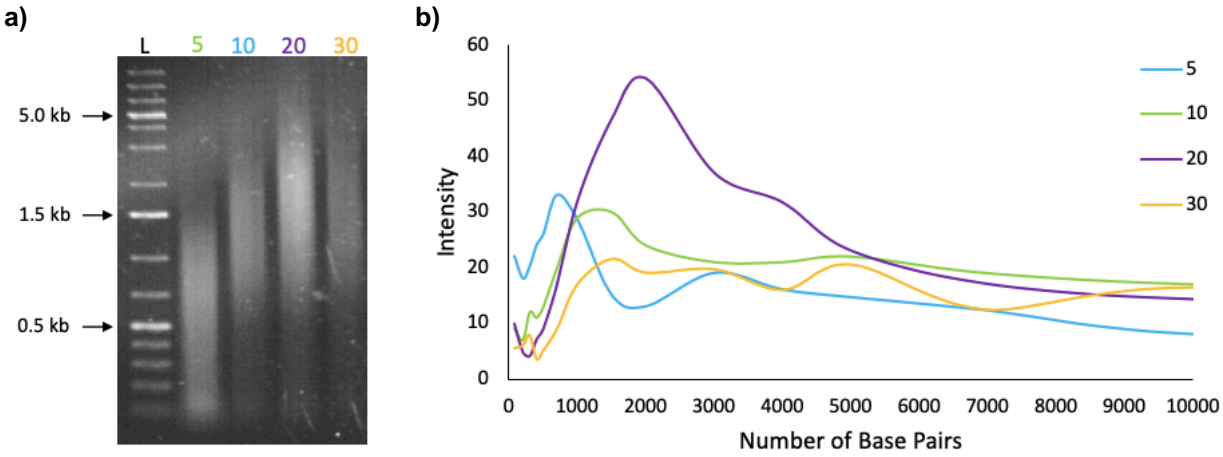


Figure 3.1. *Whitfield enzymatic extension with Tgo-pol Z3 exo⁻ polymerase after 5, 10, 20 and 30 cycles.* ***a)*** *1 % agarose gel depicting product length after each heat-cool cycle set. L=DNA ladder.* ***b)*** *Image J analysis of the average and the range of DNA lengths after each heat-cool cycle set.*

The 1 % agarose gel, **Figure 3.1a**, confirmed that the D base set was incorporated into the sequence to produce ABCD-DNA as the oligoseed extended further in length with the increase in heat-cool cycle lengths up to 20 cycles. 30 heat-cool cycles extended to similar lengths as 20 cycles, which was evident from both the agarose gel and the Image J analysis. The increase in extDNA length from 5 to 20 heat-cool cycles was also evident from the Image J graph, **Figure 3.1b,** as the peaks present shifted towards higher base pair lengths. The average extDNA length for 5, 10, and 20 heat-cool cycles were 700, 1,500 and 2,000 bp, respectively. 20 heat-cool cycles were determined to be the optimum condition to generate long base pair length DNA, as more heat-cool cycles, such as 30, offered no length advantage. The concentration of the extDNA produced was calculated from the UV-Vis spectrum, as the absorbance value at 260 nm was indicative of DNA. The absorbance values for the Whitfield extension of DNA-1 followed the expected pattern of increasing concentration value with respect to increasing number of heat-cool cycles, **Table 3.2**.

Table 3.2. *UV-Vis absorbance data recorded at 260 nm and calculated concentration values.*

No. of Cycles	Abs @ 260 nm	Concentration (ng/µL)
5	1.33	66.5
10	2.68	134.0
20	2.82	141.0
30	5.84	292.0

Although it was previously determined from **Figure 3.1** that 20 heat-cool cycles were the optimum conditions for extDNA production, it was clear from the concentration values that 30 heat-cool cycles produced a higher quantity of DNA. This was expected as the DNA would continue to replicate and grow, however since the length from the agarose gel showed no advantage over the 20 cycles, it was deemed that the concentration and length from 20 cycles was sufficient. In order to ensure that the Whitfield extension of DNA-1 had occurred by the incorporation of dG bases and not by base mismatch, snake venom digestion of the extDNA sample was performed overnight at 37 °C. 50 µL of the digested sample was injected into the HPLC for analysis and was compared to three controls; dNTP standards, oligo-1 (5'-$A_9GC_{10}T_9C$-3') and oligo-2 (5'-$CT_9C_{10}GA_9$-3'), **Figure 3.2**.

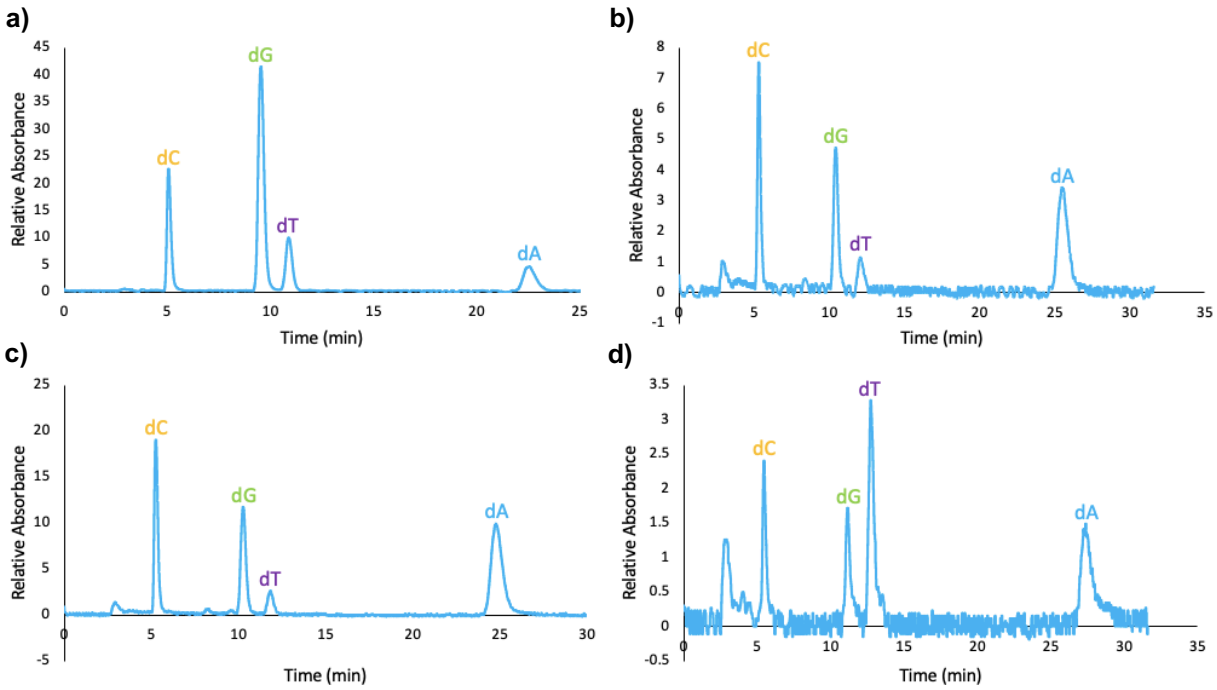


Figure 3.2. *HPLC traces of digested ABCD-DNA after enzymatic extension* ***a)*** *dNTP standards* ***b)*** *Oligo-1: 5'-$A_9GC_{10}T_9C$ -3' c) Oligo-2: 5'-$CT_9C_{10}GA_9$ -3' d) Enzymatically extended ABCD-DNA.*

The dNTP standards, **Figure 3.2a**, showed the retention time, **Table 3.3**, of the bases in the sampling conditions (70 % buffer C and 30 % buffer B for 25 minutes). This indicated the expected retention time at which the bases should elute in the oligo controls and the digested extDNA samples. It was clear from **Figure 3.2b and c**, oligo-1 and oligo-2, that dG base was present in the starting oligo seed, observed by the dG peak at 10.5 and 10.3 minutes. After extension, **Figure 3.2d**, it was evident that the dG peak in the HPLC chromatogram eluted at 11.2 minutes.

Table 3.3. *HPLC retention times for dNTP standards and ABCD-DNA, before and after enzymatic extension.*

Base	Retention Time (mins)			
	Standards	Oligo-1	Oligo-2	ABCD
dC	5.4	5.3	5.3	5.4
dG	10.5	10.5	10.3	11.2
dT	12.3	12.3	12.0	12.7
dA	25.6	25.6	24.8	27.3

The HPLC analysis failed to characterise the incorporation of the D base set, G bases, into the ABCD-DNA as the starting oligo contained dG. Therefore, no new base identification could be observed in the HPLC chromatogram to determine correct base incorporation. A new oligoseed which did not contain all bases would need to be designed to observe the new base incorporation.

3.2.2 DNA-2 $[(A_4T)_2C_{10}(T_4A)_2]$

The DNA-2 oligoseed was designed and contained only three bases in the A, B and C sections of the sequence, $[(A_4T)_2C_{10}(T_4A)_2]/[(A_4T)_2C_{10}(T_4A)_2]$, allowing for section D to be easily characterised by HPLC. For the insertion of the D section into the ABCD-DNA, Whitfield enzymatic extension was performed for 5, 10, 15 and 20 heat-cool cycles with *Thermococcus gorgonarius* polymerase B (Pfu-Pol) Z3 exo$^-$, **Figure 3.3**.

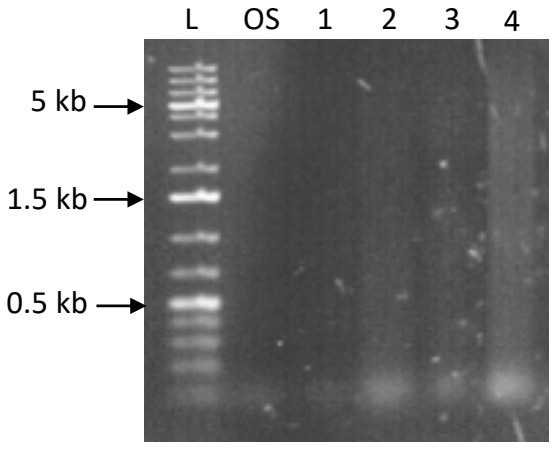


Figure 3.3. *1 % agarose gel of DNA-2 after 5, 10, 15 and 20 enzymatic extension cycles.*

The Whitfield enzymatic extension was not successful for the DNA-2 oligoseed as there was no growth in DNA length observed in the 1 % agarose gel. The 260 nm absorbance data also reflected the failed enzymatic extension of the DNA-2 oligoseed, **Table 3.4**. The

concentrations produced were much lower than expected for extDNA, especially after 20 heat-cool cycles. From previous research[18], it was expected that the 20 heat-cool cycles would generate DNA concentrations of 50-100 ng/µL depending on the sequence design.

Table 3.4. *Absorbance at 260 nm for DNA-2 enzymatic extension products after 5, 10,15 and 20 cycles.*

No. of Cycles	Abs @ 260 nm	Concentration (ng/µL)
5	0.16	8.2
10	0.26	13.2
15	0.25	12.3
20	0.34	17.2

As the Whitfield enzymatic extension of DNA-1 was successful, the low concentration values indicated that the DNA-2 oligoseed was difficult to enzymatically extend due to its complicated repeat design. The extension parameters would need to be changed to generate longer length extDNA-2. The annealing step is one of the most important steps as it determines whether the DNA will extend or not depending on its formation once annealed, so the melting temperature (Tm) of DNA-2 was studied. This allowed for a better understanding of the annealed DNA species present and indicated the best annealing temperature *i.e.* 5 °C below that of the Tm. The Tm was determined by running a UV-Vis temperature scan from ramping the temperature of the UV-Vis chamber from 15 °C to 95 °C, **Figure 3.4**.

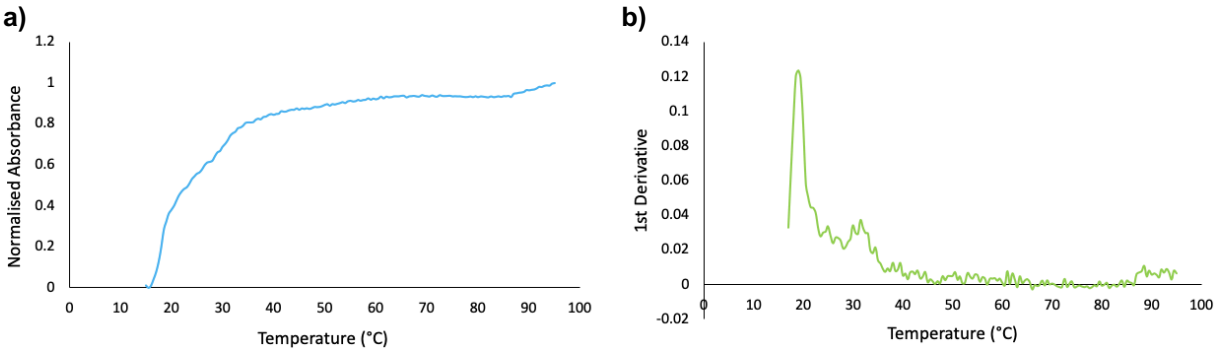


Figure 3.4*. Melting temperature of DNA-2.* ***a)*** *Melting curve* ***b)*** *1st derivative.*

The melting curve, **Figure 3.4a**, showed hyperchromicity[19] as there was an increase in UV-Vis absorbance caused by the denaturation of the dsDNA into two single strands. The shape of the DNA-2 melting curve was not the ideal S-shape observed in DNA melting as it was evident that there were two melting transitions. On annealing, before Tm analysis,

the oligoseed could form two duplex structures caused by the repetitive nature of the A base set, $(A_4T)_2$, and the C base set, $(T_4C)_2$, **Scheme 3.9**.

Full annealing of A' to C:

5'-AAAATAAAATCCCCCCCCCCTTTTATTTTA-3'
3'-AAAATAAAATCCCCCCCCCCTTTTATTTTA-5'

Partial annealing of A' to C:

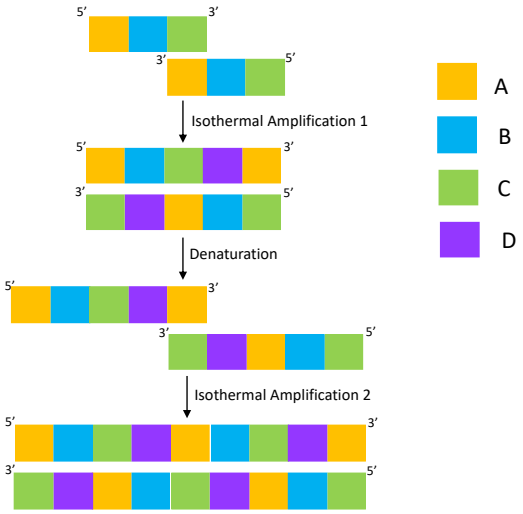
5'-AAAATAAAATCCCCCCCCCCTTTTATTTTA-3'
3'-AAAATAAAATCCCCCCCCCCTTTTATTTTA-5'

Scheme 3.9. *Potential duplex structures formed during DNA-2 oligoseed annealing after 10 minutes at 95℃ and cooling to room temperature.*

Two Tm peaks were clear from the first derivative, **Figure 3.4b**, where it was evident that there was a main Tm at 19.5 °C and a second Tm observed at 17.5 °C. These temperatures were much lower than the annealing temperature (55 °C) used in the Whitfield enzymatic extension, so poor extension could be explained by not using the optimum annealing temperature for DNA-2 formation. The Whitfield extension annealing temperature was lowered to 45 °C for 30 seconds and all other parameters kept constant. The 1 % agarose showed no evidence of DNA extension or of the presence of DNA in the UV-Vis spectra after PCR clean up purification. In case the oligoseed required more time to form the more stable duplex form during the annealing step, the temperature was kept at 45 °C and the time was increased to 60 seconds. After Whitfield extension and purification, the 1 % agarose gel and UV-Vis spectra showed no evidence of extension of the oligoseed. Finally, in addition to a decreased annealing temperature and an increase in annealing time, the concentration of the oligoseed was increased to encourage duplex formation with an excess of DNA. However, once again there was no observed extension of the oligoseed and so no insertion of the new base set, D (G_{10}).

A new approach to the extension of the DNA-2 oligoseed was investigated by filling in the oligoseed to make a more stable duplex formation for use during the Whitfield enzymatic extension. Isothermal extension of the oligoseed was performed for 1 hour at 65 °C using Bst 2.0 WarmStart DNA polymerase (NEB, UK), a homologue of *Bacillus stearothermophilus* DNA polymerase I. The isothermal extension would allow for the DNA-2 oligoseed to anneal in the 3'-ABC-5'/5'-ABC-3' formation which permits the

polymerase to fill in the sticky ended overhang and form a more stable duplex with a higher Tm, **Scheme 3.10**.



Scheme 3.10. *Isothermal extension of 3'-ABC-5'/5'-ABC-3' oligoseed using Bst 2.0 WarmStart polymerase.*

The components of the isothermal extension reaction were added to a 0.6 mL Eppendorf, excluding the polymerase and the samples were placed in the PCR, to allow for temperature-controlled reactions. Initially, the DNA-2 oligoseed was treated to 95 °C for 2 minutes to fully denature any duplex structures formed at room temperature and then cooled to 25 °C for controlled re-annealing. The Bst 2.0 polymerase was added and the samples held at 65 °C for 1 hour which enabled isothermal extension of the oligoseed, isothermal amplification 1. The sample was heated to 95 °C for 2 minutes and cooled to 25 °C for denaturation of the new oligoseed duplex structure and re-annealing to occur. In the final stage, isothermal amplification 2, another aliquot of Bst 2.0 polymerase was added and the samples heated to 65 °C for 1 hour before they were cooled to room temperature. By performing 2 isothermal extension cycles, the oligoseed should have increased in base pair length to a more stable duplex structure, up to 90 bp in length. As a result, the annealed overlap would be more stable during the Whitfield enzymatic extension and therefore more likely to survive extension before denaturing due to a low Tm.

After isothermal extension, the samples were not purified as they would not have increased enough in length to bind to the purification column. This was due to the limitations of the column, as any DNA < 100 bp would elute through the column during

the binding step. The reaction components, not already present in the Eppendorf from the isothermal extension, were added to the Eppendorf (10x reaction buffer, dNTPs and Z3 polymerase) to make a final volume of 50 μL. Whitfield extension was performed for 20 heat-cool cycles and purified using a PCR clean up kit before characterisation by agarose gel electrophoresis, **Figure 3.5**.

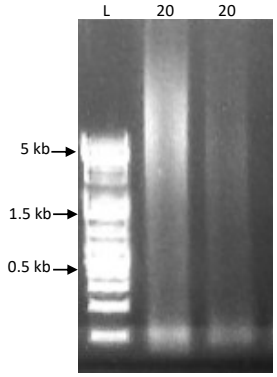


Figure 3.5. *1 % agarose gel of DNA-2 after enzymatic extension 2 isothermal cycles followed by the Whitfield enzymatic extension method for 20 heat-cool cycles.*

This time the 1 % agarose gel indicated that although there was some very short DNA at approximately 75 bp was present, the DNA-2 oligoseed, after isothermal and Whitfield extension, had extended in length to over 20,000 bp. In comparison to the previous Whitfield extension attempt with DNA-2 oligoseed, **Figure 3.3**, this attempt yielded a longer base pair length product and produced a higher concentration of DNA after PCR clean up kit purification, **Table 3.5**. The concentrations of the two extended samples varied as one sample produced more concentrated extension products. This was also evident in the agarose gel as lane 1 was brighter than lane 2, however, the concentration values here were much higher than those achieved in **Table 3.4**. This further suggested that D (G_{10}) had been successfully inserted into the ABC sequence during the isothermal extension as the more stable oligoseed grew much longer in length during the Whitfield extension than the starting oligoseed.

Table 3.5. *Absorbance at 260 nm for DNA-2 isothermal and enzymatic extension.*

No. of Cycles	Abs @ 260 nm	Concentration (ng/μL)
20	3.89	194.3
20	1.42	71.2

For more accurate determination of the extDNA base pair length, AFM imaging was employed to visualise the DNA. The DNA-2 20 heat-cool cycle extended sample was applied to a freshly cleaved mica surface at 4 ng/µL with 0.5 mM $MgCl_2$ to immobilise the DNA to the surface, **Figure 3.6**.

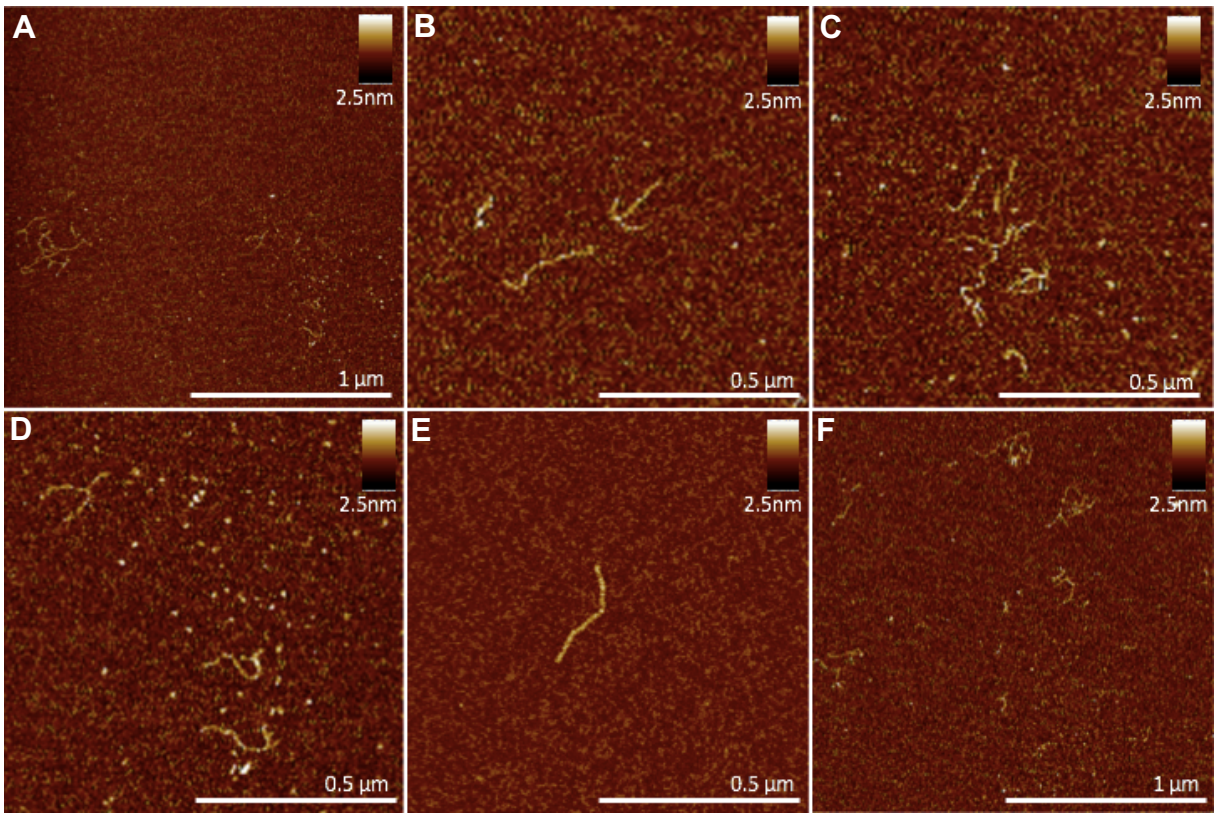


Figure 3.6. *AFM of DNA-2 after isothermal and 20 heat-cool enzymatic extension cycles. Molecularly combed 4 ng/µL DNA with 0.5 mM $MgCl_2$ on a freshly cleave mica surface.*

The AFM images **A-F** showed that only a minimal amount of DNA remained on the mica after the wash step during surface preparation. It was evident that there were single strands of dsDNA in addition to and networked structures present in the sample. Image **D** also highlighted the presence of very short DNA (< 200 bp), which showed as dots on the image. Due to the folded and networked nature of the DNA imaged, only a few strands could be measured to obtain an average length of 379 ± 164 bp (129 ± 56 nm). The average height of the extDNA-2 was 1 ± 0.2 nm. The measured length confirmed that the extDNA-2 had elongated to lengths longer than the oligoseed and thus incorporation of the D base set was successful. To further characterise the successful incorporation, the Whitfield extension was performed for 5 heat-cool cycles after isothermal extension to generate a sample which was longer than the starting oligoseed but short enough (< 500 bp) to send to GATC biotech for sequencing analysis to be run. The ABC primer,

$(A_4T)_2C_{10}(T_4C)_2$, was sent alongside the ABCD-extDNA-2 for the sequence to be analysed, **Figure 3.7**.

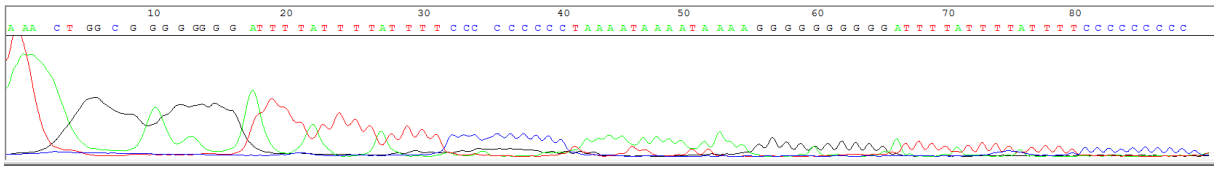
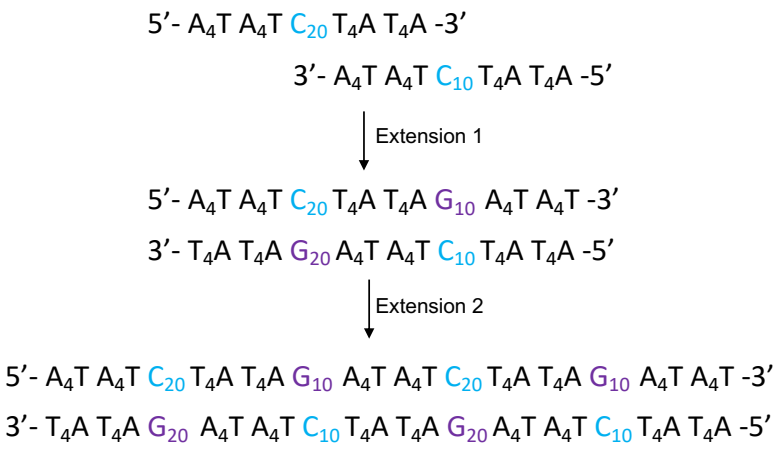


Figure 3.7. *DNA-2 Sequencing by GATC Biotech.*

Although the start of the sequencing did not match correctly to the expected sequence, the sequencing data showed that the oligoseed had extended and incorporated the correct number and type of bases. Therefore, it was concluded that successful incorporation of the bases contained in section D was achieved and the completed DNA sequence was ABCD. By being able to add the D section into the sequence, a run of 10 dG bases were incorporated in a timely manner, which would have potentially taken weeks to synthesise by automated synthesis. This proved that DNA with difficult repeat units could be easily synthesised by enzymatic methods, giving users the ability to design a wider variety of sequences for specific applications.

3.2.3 DNA-3 $[(A_4T)_2C_{20}(T_4A)_2]$

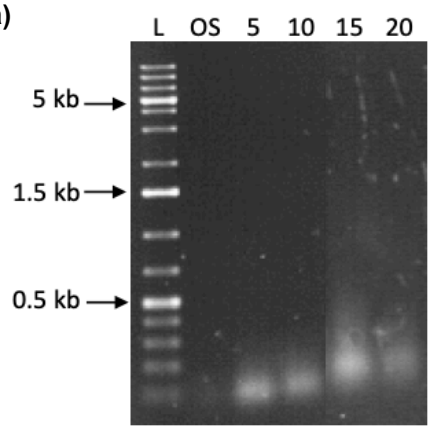
Sections 3.2.1 and **3.2.2** proved that the ABCD-DNA method was successful in incorporating the new sequence section D into the designed ABC oligoseed. This generated long DNA base repeats at a pre-designed spatial distance which could be used as recognition binding sites. In order to accommodate multiple sizes of binding sites within the one DNA sequence, the $(A_4T)_2C_{10}(T_4A)_2$ oligo from DNA-2 was paired with $(A_4T)_2C_{20}(T_4A)_2$ to form DNA-3. From the extension of DNA-3, the D section of the ABCD-DNA structure would alternate between G_{10} and G_{20} every 10 bases, **Scheme 3. 11**.



Scheme 3.11. *ABCD-DNA structure of DNA-3, with alternate length binding sites in section D.*

The DNA-3 oligoseed underwent extension by the Whitfield enzymatic method for 5, 10, 15 and 20 heat-cool cycles using *Thermococcus gorgonarius* polymerase B (Pfu-Pol) Z3 exo$^-$. The products were purified using a PCR clean up kit and analysed by agarose gel electrophoresis and UV-Vis spectroscopy, **Figure 3.8**.

a)



b)

No. of Cycles	Abs @ 260 nm	Concentration (ng/μL)
5	1.33	66.4
10	2.68	133.8
15	2.82	141.0
20	5.84	292.0

Figure 3.8. a) *1 % agarose gel of DNA-3 after 5, 10, 15 and 20 Whitfield enzymatic extension heat-cool cycles.* ***b)*** *UV-Vis absorbance at 260 nm and concentration values.*

It was clear from the 1 % agarose gel that some extension of the DNA-3 oligoseed had occurred since the DNA in each lane of the gel was longer than that of the initial length, 50 bp. If taken in isolation, the absorbance data, **Figure 3.8b**, would suggest that the DNA extension was successful in a similar manner as previous oligoseed extensions due to the high (> 20 ng/μL) and increasing concentration values for each heat-cool cycle length. However, as identified by the agarose gel, it was observed that DNA-3 generated high concentrations of short DNA lengths, especially for 20 heat-cool cycles as the concentration was 292 ng/μL. For all heat-cool cycle lengths, the DNA was < 500 bp and therefore, unsuitable for spatial binding studies. It was believed that the short DNA

lengths produced may have been caused by the unstable duplex structure due to the potential formation of two oligoseed structures, as identified with the DNA-2 oligoseed (**Section 3.2.2**) or the variation in G length in the D sequence of the ABCD-DNA format. To gain a better understanding of the annealing process of the oligoseed, Tm analysis was performed on a UV-Vis spectrometer with temperature chamber and swept from 5 °C to 95 °C and the absorbance recorded at 260 nm, **Figure 3.9**.

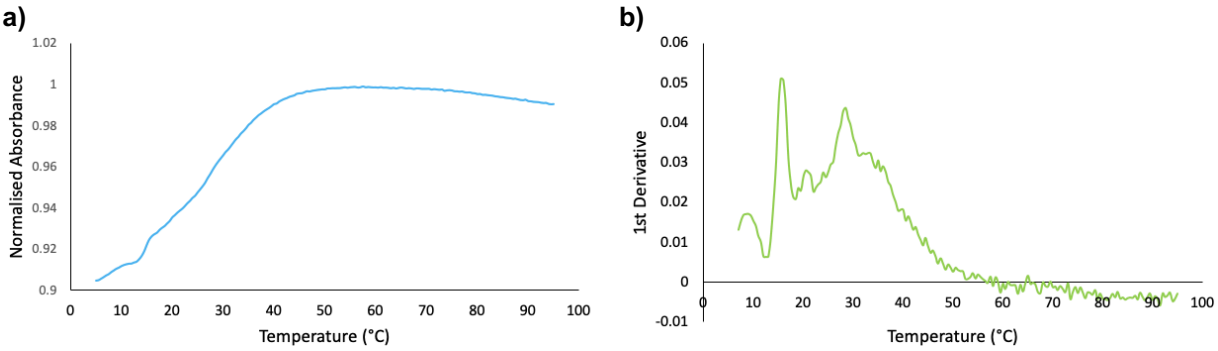


Figure 3.9. *Melting temperature of DNA-3 oligoseed.* ***a)*** *Melting curve* ***b)*** *1st derivative.*

The DNA-3 oligoseed melting curve showed that hyperchromicity was present in the sample as the UV-Vis absorbance values increased as the temperature increased, hence dsDNA melting to ssDNA. The 1st derivative had many melting peaks evident and showed that the dsDNA-3 oligoseed contained various annealed formations, all of which had lower Tms (10 , 16 , 21.5 and 29 °C) than the annealing temperature in the Whitfield method, 55 °C. In order to try and stabilise the oligoseed duplex structure, 2 mM $MgSO_4$ was added to the DNA-3 solution and the temperature ramp repeated under the same sampling conditions, **Figure 3.10**.

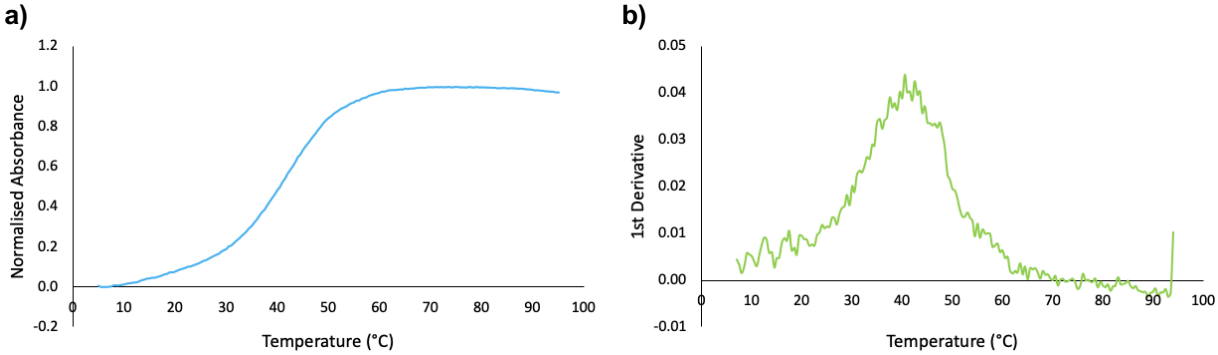
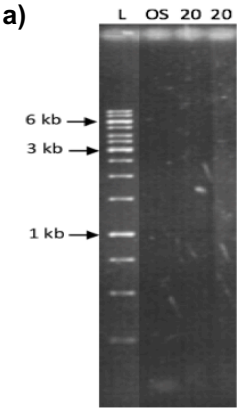


Figure 3.10. *Melting temperature of DNA-3 oligoseed with* $MgSO_4$. ***a)*** *Melting curve* ***b)*** *1st derivative.*

With the added $MgSO_4$ salt, the shape of the melting curve improved which suggested that there were fewer melting episodes occurring and therefore, fewer dsDNA-3 oligoseed structure variations. The 1st derivative showed the stabilisation of the oligoseed with the

added salt as a single peak was present with a Tm of 42.5 °C. Subsequently, the annealing step in the Whitfield enzymatic extension was lowered from 55 °C to 40 °C and the isothermal extension was performed before the Whitfield extension. Two isothermal extension cycles using Bst 2.0 WarmStart DNA polymerase were run before the Whitfield extension was performed for 20 heat-cool cycles with added $MgSO_4$ for oligoseed stabilisation. The extension products were characterised by agarose gel electrophoresis and UV-Vis spectroscopy, **Figure 3.11**.

a)



b)

No. of Cycles	Abs @ 260 nm	Concentration (ng/µL)
20	1.20	60.2
20	0.45	22.3

Figure 3.11. a) *1 % agarose gel of DNA-3 after isothermal and enzymatic extension after 20 heat-cool cycles* ***b)*** *UV-Vis absorbance at 260 nm and concentration values.*

The 1 % agarose gel showed minimal evidence of DNA-3 extension compared to the oligoseed in lane OS. The UV-Vis absorbance data at 260 nm proved that some extension may have occurred as the concentrations of the samples were above 20 ng/µL, however the DNA produced must be very short as it was not observed on the agarose gel. It was therefore concluded that the extension of the DNA-3 ABCD-DNA sequence was poor due to the difficult sequence design. As the DNA was very short and undetectable, the sequence could not be used for spatial binding studies.

3.2.4 DNA-4 [$C_{10}A_{10}G_{10}$]

[$C_{10}A_{10}T_{10}$]/[$C_{10}A_{10}T_{10}$] (DNA-4) was a sequence containing long poly repeats of bases and was designed for ease of DNA annealing and characterisation to produce DNA suitable for use in spatial binding studies. Initially, the Tm of the short 3'-ABC-5'/5'-ABC-3' oligoseed was determined by UV-Vis thermal scanning, **Figure 3.12**.

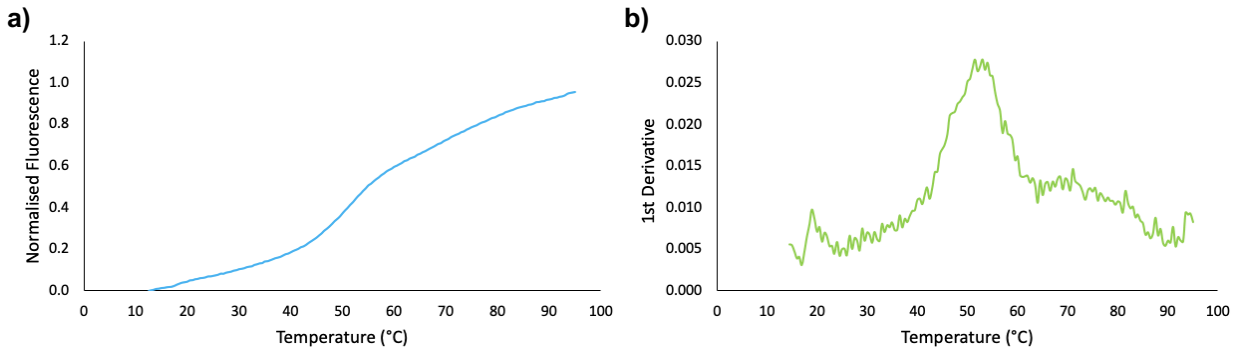


Figure 3.12. *Melting temperature of DNA-4 oligoseed.* ***a)*** *Melting curve* ***b)*** *1st derivative.*

As discussed in **Section 3.2.2**, the melting curve had the expected shape for DNA melting and showed that melting of the duplex had occurred. The first derivative revealed that the Tm of the duplex was 52 °C, which was higher than the predicted Tm of 40 °C from OligoCalc. Since the Tm was much higher for this oligoseed than the others which were previously discussed in this chapter, no isothermal extension was performed. The oligoseed was extended by only using the Whitfield enzymatic extension method and characterised by agarose gel electrophoresis and Image J analysis, **Figure 3.13**.

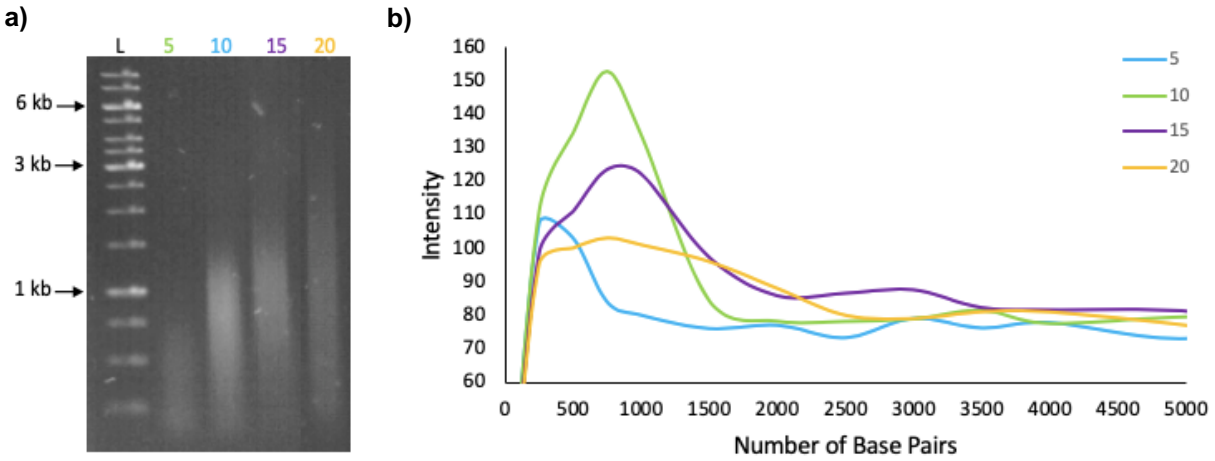


Figure 3.13. *Enzymatic extension with Tgo-pol Z3 exo (-) polymerase after 5, 10,15 and 20 cycles.* ***a)*** *Agarose gel depicting each cycle set. L=DNA ladder.* ***b)*** *Image J analysis of DNA length after each cycle set.*

The Whitfield enzymatic extension of DNA-4 was carried out for 5, 10, 15 and 20 heat-cool cycles with *Thermococcus gorgonarius* polymerase B (Pfu-Pol) Z3 exo$^-$. The 1 % agarose gel, **Figure 3.13a**, showed good extension of DNA-4 without the use of the isothermal extension of the oligoseed prior to PCR cycling. Furthermore, step-wise growth was evident as the average length increased for each of the heat-cool cycle lengths. This finding was in accordance with the Image J plot profile, **Figure 3.13b**, as the peak maximum (average base pair length) increased from 500, 1,000, 1,200 to 1,500 bp, for

5,10, 15 and 20 heat-cool cycles, respectively. Further characterisation of the PCR product was undertaken by UV-Vis spectroscopy, **Table 3.6**.

Table 3.6. *UV-Vis absorbance recorded at 260 nm and calculated concentration values.*

No. of Cycles	Abs @ 260 nm	Concentration (ng/μL)
5	1.03	51.6
10	1.82	90.9
15	2.56	127.9
20	2.82	140.9

As expected, the absorbance values at 260 nm showed an increase in concentration as the heat-cool cycle lengths increased. The concentration values depicted the success of the Whitfield enzymatic extension for the DNA-4 oligoseed due to the high concentrations observed, particularly after 15 cycles. To confirm the base pair length with greater accuracy and to visualise the DNA, AFM imaging was employed by depositing 4 ng/μL of 20 heat-cool cycle extDNA-4 on a freshly cleaved mica surface with 0.5 mM $MgCl_2$, **Figure 3.14**.

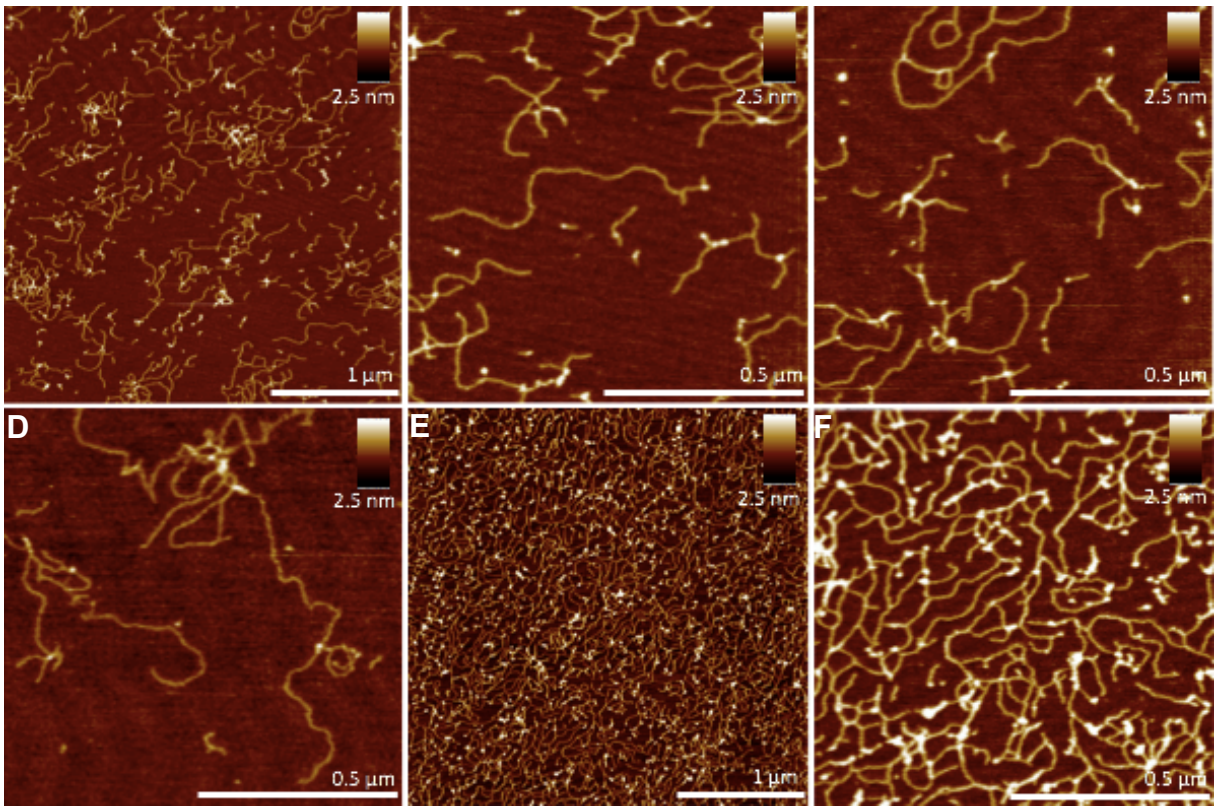


***Figure 3.14**. AFM images of ABCD-DNA after Whitfield extension for 20 heat-cool cycles at 4 ng/µL on freshly cleaved mica with 0.5 mM $MgCl_2$.*

It was clearly evident that the DNA-4 3'-ABC-5'/5'-ABC-3' oligoseed had successfully formed ABCD-DNA and extended well in base pair length to generate long length DNA which contained four different long repeat sections. Various extDNA-4 structures were observed in the AFM images as indicated in **A** which contained highly networked, self-annealed and folded structures, along with molecularly combed single dsDNA strands. It also captured the base pair size distribution evident in the agarose gel (**Figure 3.13**), since many different lengths can be seen in the 3 µm images (**A** and **C**). **B-D** showed some of the extDNA-4 as singular double strands of DNA which allowed for easy measurement of the base pair length, although the lengths of these strands were shorter than would be expected of 20 heat-cool cycle extDNA. The longer length DNA tended to be networked or folded and not accurately measurable. However, measuring these strands gave an indication of the average base pair length which was was 706 ± 465 bp (240 ± 158 nm). Also, images **B-D** showed the more intricate detail of the looped and networked structures of the very long DNA. This DNA could be formed from one strand of DNA or it could be formed of several DNA strands resulting in concatemeric structures which have annealed to form the detailed patterns.

The average height of the extDNA-4 was 0.8 ± 0.2 nm and was comparable to the dsDNA literature value of 0.5 – 1.0 nm.[20] Moreover, images **E** and **F** exhibited the DNA networks which extDNA-4 formed at the higher end of the concentration gradient on the mica generated by molecular combing. It was clear that some of the network structures contained DNA which was folded or self-annealed within the structure and this was reflected in the elevated average height value of 1.7 ± 0.7 nm and also, in the difference of the visual thickness of the DNA strands which comprised the network. Despite the various structures present, the AFM images further confirmed good extension of the DNA-4 oligoseed due to the large increase in DNA length which was initially 30 bases, with a 10 bp overlap before Whitfield extension.

In order to determine whether the correct ABCD-DNA sequence was enzymatically synthesised, *i.e.* the correct bases were inserted by the polymerase during the heat-cool cycles, a 5 heat-cool cycle DNA-4 sample was sent to GATC Biotech for sequencing. Unfortunately, the sequencing of this ABCD product failed due to the long repetitive nature of the sequence, which rendered it difficult for the machine to accurately read. Therefore, to ensure the ABCD sequence contained the correct bases and to further confirm the successful insertion of the T bases in the new D section of the DNA, HPLC analysis was employed after snake venom digestion of the single stranded oligos and ABCD-DNA, **Figure 3.15**.

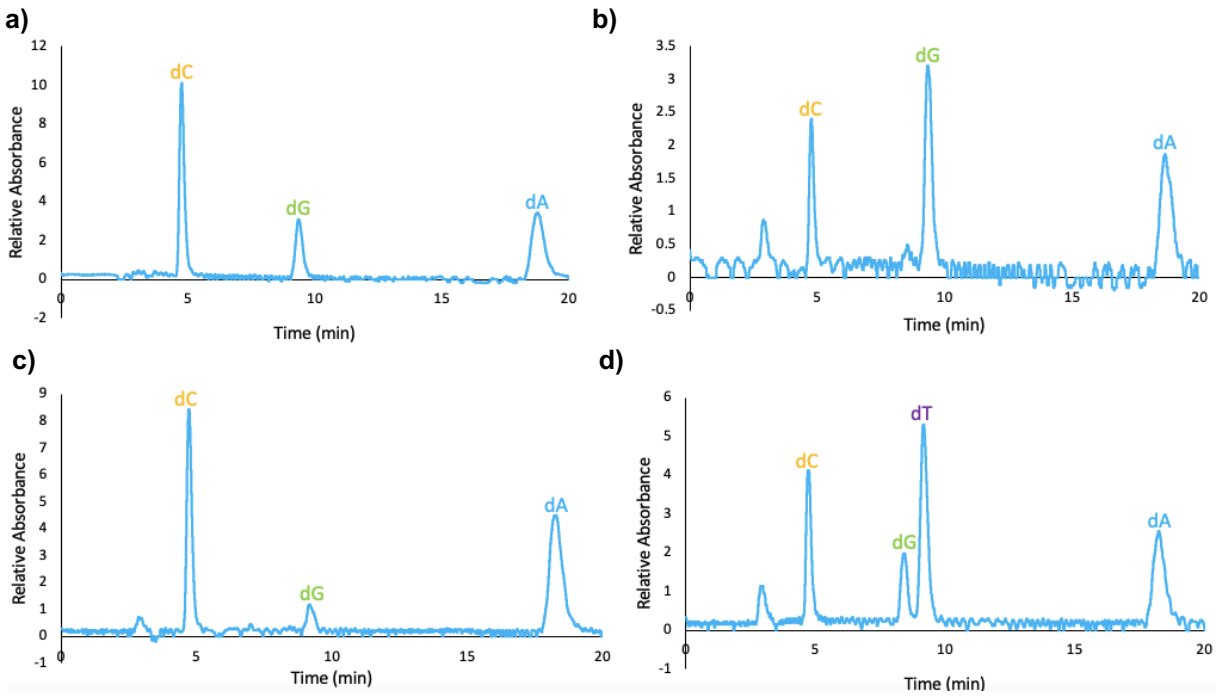


Figure 3.15. *HPLC traces of digested ABCD-DNA after enzymatic extension* ***a)*** *dNTP standards* ***b)*** *Oligo-1: 5'-$C_{10}A_{10}T_{10}$-3'* ***c)*** *Oligo-2: 5'-$G_{10}A_{10}C_{10}$-3'* ***d)*** *Enzymatically extended ABCD-DNA.*

The HPLC system was run with sampling conditions of 70 % buffer C and 30 % buffer B for 25 minutes. Initially 1 mM dNTP standards, without dTTP, were injected into the HPLC to determine the standard elution times at which to expect the commercially synthesised oligoseeds, oligo-1 (5'-$C_{10}A_{10}G_{10}$-3') and oligo-2 (5'-$G_{10}A_{10}C_{10}$-3'), **Figure 3.15b** and **c**, respectively. Oligo-1 and oligo-2 showed that there was no dT bases present in the starting oligoseed as there was no corresponding HPLC peak present in any of the traces. Furthermore, the retention times of oligo-1 and oligo-2, **Table 3.7**, agreed well with the retention times of the dNTP bases which confirmed the bases were correct and as expected.

The digested ABCD-DNA sample, **Figure 3.15d**, was injected and it was clear that four peaks had eluted from the column, rather than just three as seen in the oligo samples. The new peak in the HPLC trace was attributed to the introduction of D into the ABC sequence which created the ABCD structure and with this insertion produced a T_{10} base section. The retention times of dC (4.7 min) and dA (18.3 min) matched well to the other sample retention times and so were easily identified. The dG bases were eluted from the column slightly earlier (8.5 min) than the oligo samples (9.2 min) due to the presence of the dT base, which forced the dG bases to elute quicker from the column. This was because the two bases had very similar retention times and in the incorrect buffer conditions would elute as one peak in the HPLC trace.

Table 3.7. *HPLC retention times for ABCD-DNA, before and after enzymatic extension.*

Base	Retention Time (mins)			
	dNTPs	Oligo-1	Oligo-2	ABCD
dC	4.8	4.7	4.7	4.7
dG	9.4	9.3	9.2	8.5
dT	n/a	n/a	n/a	9.2
dA	18.9	18.7	18.3	18.3

Overall, from the characterisation of the ABCD-DNA product, it was evident that DNA-4 incorporated the new base set D successfully and extended to base pair lengths which were expected from the Whitfield extension method, increasing length with increasing heat-cool cycles. Furthermore, from the insertion of the repeat poly(T) sequence in the new D section of the DNA, DNA-4 was a good candidate to incorporate long repeats of

modified bases which are otherwise difficult to include when designing an oligo for automated synthesis.

3.2.5 DNA-4 with Modifications

As discussed in **Section 2.1**, the synthesis of modified base DNA oligos to generate spatial binding sites within the DNA sequence can be difficult by automated synthesisers and other enzymatic methods, RCA, slippage *etc.* The sequence design is generally limited to having only a few of the same bases next to one another for automated synthesis and to commercially available DNA sequences for enzymatic synthesis methods. Generating designer sequences by the ABCD-DNA method would allow for more user-defined sequence design and larger areas for proteins or molecules to bind to. The ABCD-DNA method allowed for ease of design and production of potential recognition sites as demonstrated in **Sections 3.2.1**-**3.2.4** with standard dNTP incorporation into oligoseeds DNA-1, DNA-2 and DNA-4. These sections showed that it was possible to enzymatically incorporate a new set of bases, termed as D, which was 10 bases in length and not present in the original oligoseed sequence, 3'-ABC-5'/5'-ABC-3'. Therefore, to generate more targeted DNA spatial recognition sites, the incorporation of modified bases was desired. The ABCD-DNA enzymatic synthesis approach was employed to incorporate modified bases as the new base section D and generate long length DNA with set distance spatial binding sites.

Phosphorothioate modified bases were used to create spatial recognition sites for AuNP binding. The DNA-4 oligoseed was used with the isothermal and Whitfield enzymatic extension to incorporate dATP-αS and dTTP-αS bases in place of the standard dNTP. This would generate DNA with a modified backbone, containing P-S bonds instead of P-O bonds.

In order to determine the capability of the DNA-4 sequence to incorporate phosphorothioate modified bases by ABCD-DNA enzymatic synthesis, dATP-αS was initially introduced in the B section of the DNA sequence instead of the standard dATP base. It was thought that inserting a modified base which was not in the new section of the DNA would be easier and would help to identify the optimum reaction conditions to obtain the modified ABCD-DNA.

Since the DNA-4 oligoseed incorporated the D section and extended well in base pair length by the Whitfield enzymatic extension method, this method was used initially for the addition of the dATP-αS base. The heat-cool cycles were performed for 5, 10, 15 and 20 cycles with *Thermococcus gorgonarius* polymerase B (Pfu-Pol) Z3 exo$^-$, however the results did not yield positive results for DNA synthesis. The concentration of the samples was low (< 20 ng/μL), and the agarose gel showed no extension of DNA-4, so no incorporation of the dATP-αS base. Instead, the isothermal extension, as per **Section 3.2.2**, **Scheme 3.10**, was used on the oligoseed before Whitfield extension was performed in order to form a more stable duplex, **Figure 3.16**.

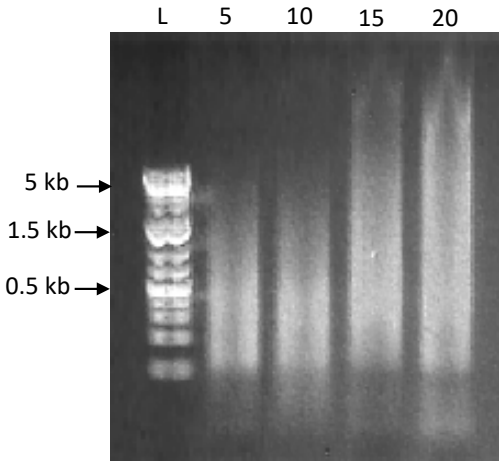


Figure 3.16. *Modified ABCD-DNA containing dATP-αS modification. 1 % agarose gel after 5, 10, 15 and 20 heat-cool cycles of Whitfield enzymatic extension.*

After isothermal and enzymatic extension, it was evident from the 1 % agarose gel that the DNA-4 oligoseed had incorporated the dATP-αS and extended in base pair length. It was apparent that 5 and 10 heat-cool cycles extended to the same length, with a maximum length of approximately 5,000 bp. 15 and 20 heat-cool cycles extended further in length but also to the same maximum base pair length which could not be identified as they had grown longer than the maximum DNA ladder length, 20,000 bp. The starting oligoseed length, < 75 bp, was evident in the agarose gel for all heat-cool cycle lengths which identified that not all of the oligoseed reacted during the heat-cool cycles. However, the length to which the DNA had grown was longer than the starting length indicating successful modified base incorporation. The UV-Vis spectrum of dATP-αS modified ABCD-DNA proved that the 260 nm peak remained unchanged after modification,

however the absorbance data did not show the typical trend of increasing concentration value relative to the number of heat-cool cycles, **Table 3.8**.

Table 3.8. *Absorbance at 260 nm for isothermal and enzymatically extended ABCD-DNA, where D was dATP-αS.*

No. of Cycles	Abs @ 260 nm	Concentration (ng/μL)
5	0.82	40.9
10	1.91	95.3
15	1.35	67.5
20	1.13	56.4

It was expected that as the number of heat-cool cycles increased, the concentration/absorbance values would also increase. This was not observed for the dATP-αS modified DNA-4 samples as the concentration increased until 10 heat-cool cycles and then started to fall for 15 and 20 heat-cool cycles. However, the concentrations obtained were consistent with successful Whitfield enzymatic extension. The concentration anomaly may have been due to the difficulty with the growth of the DNA and showed that even by using a polymerase to perform the incorporation, it was still difficult to achieve. In order to determine whether the modified DNA compared characteristically to standard DNA, the 20 heat-cool cycle sample was imaged by AFM using a 4 ng/μL sample with 0.5 mM $MgCl_2$ on freshly cleaved mica, **Figure 3.17**.

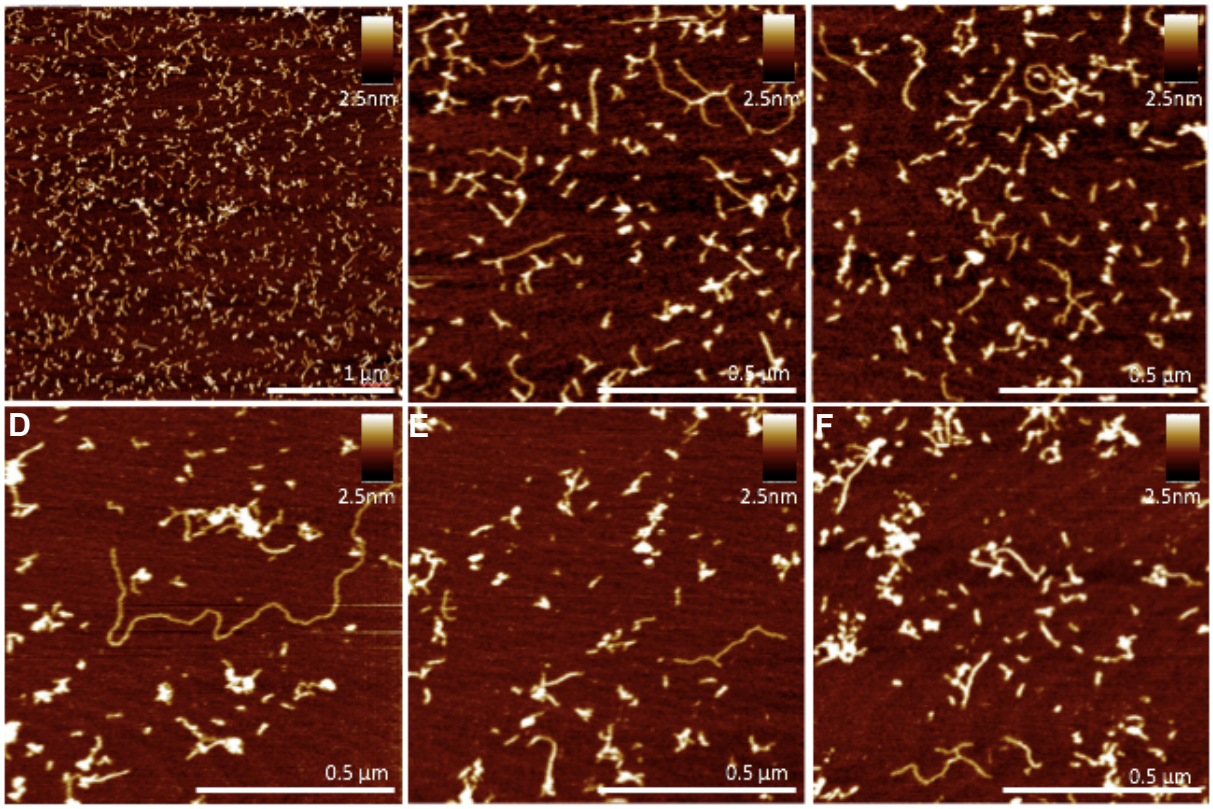


Figure 3.17. *AFM of Modified ABCD-DNA containing dATP-αS after 20 cycles enzymatic extension on freshly cleaved mica a concentration of 4 ng/µL.*

The AFM images **A**-**F** showed that the dATP-αS modified ABCD-DNA-4 had a very large size distribution as small base pair lengths (90 bp) of DNA could be observed, alongside long base pair length (1,720 bp), some of which had folded or self-annealed. Image **A** showed a 3 µm section of the mica, where it was evident that the DNA seemed to be much shorter than that observed for the unmodified ABCD-DNA-4, **Figure 3.14**. It appeared that there were more folded and self-annealed structures for this type of DNA, which may be due to the structural change in the backbone. **B**-**F** illustrated the DNA in more detail and allowed for measurement of the individual dsDNA strands. The average base pair length was 390 ± 280 bp (130 ± 100 nm) and therefore, shorter than that quoted from the agarose gel (**Figure 3.16**) due to the folded structures present in the images. The height of the dATP-αS modified ABCD-DNA-4 was 1.0 ± 0.3 nm and corresponded well to the literature values, 0.5 – 1.0 nm.[20] This confirmed that the modified ABCD-DNA-4 generated compared well characteristically to unmodified ABCD-DNA-4 as visually it appeared as standard DNA. Furthermore, it was evident that in **A**-**F**, there were some brighter areas on the modified ABCD-DNA-4 and this was attributed to the self-annealing and folding of the DNA structure due to the self-complementarity of the designed sequence. The height of these areas alluded that the DNA structure was different to that of the dsDNA structure

as the average height was measured as 2.7 ± 0.6 nm. This was a difference of 1.7 nm and confirmed the structure contained more than two DNA strands.

Since the characterisation of dATP-αS modified DNA-4 showed incorporation of the modified base and extension of the base pair length from the starting oligoseed length, the modification of the new base section D of the ABCD structure was undertaken. Isothermal and Whitfield enzymatic extension was performed on the DNA-4 oligoseed with the addition of dTTP-αS instead of the standard dNTP, dTTP. The isothermal extension was performed for 2 cycles at 65 °C for 1 hour with Bst 2.0 WarmStart polymerase and the Whitfield extension undertaken in the standard heat-cool cycle conditions with *Thermococcus gorgonarius* polymerase B (Pfu-Pol) Z3 exo$^-$ for 5, 10, 15 and 20 heat-cool cycles, **Figure 3.18**.

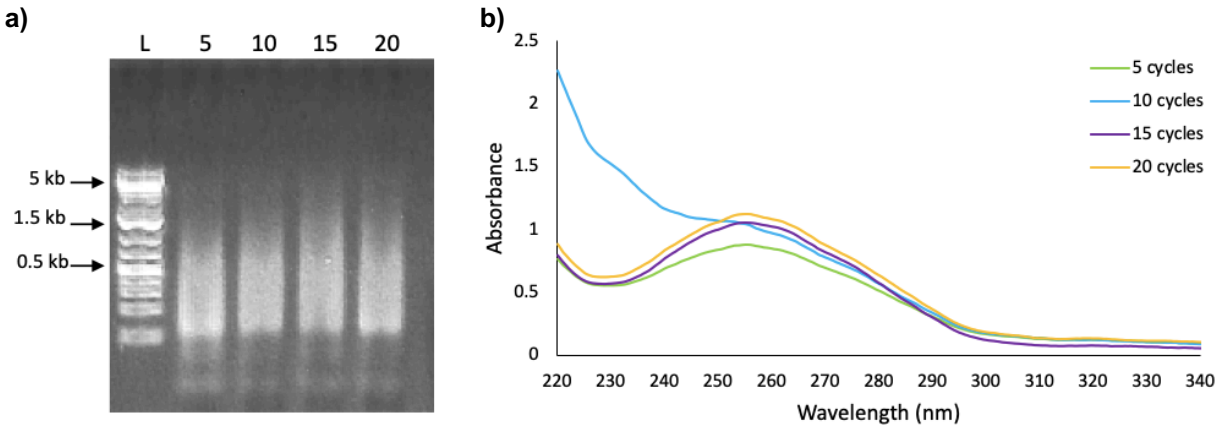


Figure 3.18. *Modified ABCD-DNA containing dTTP-αS modification.* ***a)*** *Agarose gel after 5, 10, 15 and 20 cycles enzymatic extension* ***b)*** *Absorbance at 260 nm from the UV-Vis spectra.*

After isothermal and enzymatic extension, it was evident from the 1 % agarose gel that the DNA-4 oligoseed had incorporated dTTP-αS and extended in base pair length due to the length distributions observed after 5, 10, 15 and 20 heat-cool cycles. Also, it was apparent that the increase in heat-cool cycle lengths did not help to extend the DNA base pair length as 5 cycles produced the same length distribution as 20 cycles. The average base pair length for each cycle was approximately 300 - 400 bp, although the DNA distribution was from < 75 - 5000 bp. The starting oligoseed length was evident in the agarose gel, < 75 bp, which was indicative that not all of the oligoseed had reacted during the PCR heat-cool cycles. However, the length to which the DNA had grown was longer than the starting length, which indicated successful modified base incorporation. The UV-Vis spectroscopy data, **Figure 3.18b**, showed that the modified ABCD-DNA followed

the intended trend of increasing absorbance value at 260 nm with the increase in the number of heat-cool cycles. The concentration values for 5, 10, 15 and 20 heat-cool cycles were 42.5, 48.5, 51.2 and 54.1 ng/μL, respectively. As carried out before with the dATP-αS modification, AFM analysis was performed on a 4ng/μL extDNA sample after 20 heat-cool cycles on freshly cleaved mica with 0.5 mM $MgCl_2$, **Figure 3.19**.

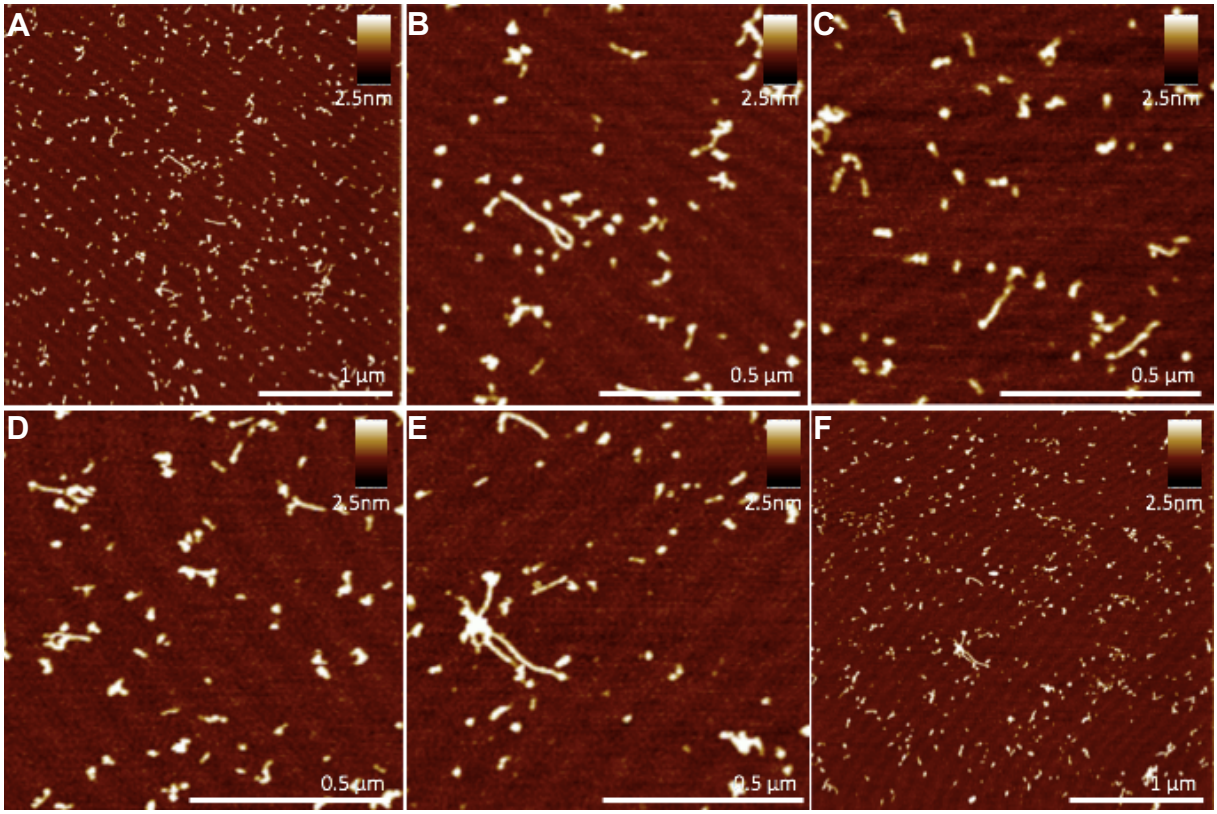


Figure 3.19. *AFM of Modified ABCD-DNA containing dTTP-αS after 20 heat-cool cycles enzymatic extension on freshly cleaved mica at a concentration of 4 ng/μL.*

The AFM images **A-F** confirmed the short nature of the DNA found in the agarose gel (**Figure 3.18**) as the average measured length of the DNA was 310 ± 180 bp (110 ± 60 nm). The average height was 1.0 ± 0.3 nm and corresponded to the literature value of 0.5 - 1 nm for dsDNA.[20] In the 3 μm images (**A** and **F**) it was clear that most of the DNA was short across the surface of the mica, however some longer DNA strands were evident in the form of looped structures and networks, which was clear in more detail in images **B** and **E**, respectively. It was also clear that some of the short DNA had self-annealed and folded back on itself, which was indicated by the brighter sections of DNA showing a difference in height from the standard dsDNA. The heights of the folded, networked and looped DNA were all similar and produced an average height of 2.3 ± 0.8 nm. Although there was a height difference across the extended dTTP-αS

modified DNA-4, it was evident that the modified base had incorporated into the ABCD structure of the sequence since the overall length was longer than that of the starting oligoseed (~ 90 bp).

Once it was clear that two different modified bases could be incorporated into the ABCD-DNA system separately, double modification of the starting 3'-ABC-5'/5'-ABC-3' oligoseed was undertaken. The DNA-4 oligoseed was used for the incorporation of dATP-αS and dTTP-αS to create a DNA sequence where the modification would be spatially separated by 10 bases on both the single strand and in the double stranded helix structure. Isothermal and Whitfield enzymatic extension was used, as per previous ABCD-DNA extension, and characterised by agarose gel electrophoresis, **Figure 3.20**.

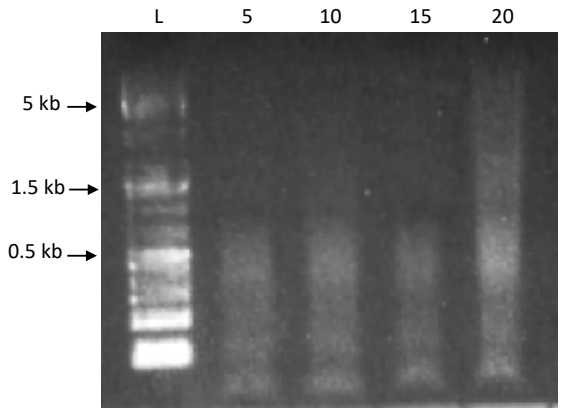


Figure 3.20. *Modified ABCD-DNA containing dATP-αS and dTTP-αS modifications. 1 % agarose gel after 5, 10, 15 and 20 cycles enzymatic extension.*

The 1 % agarose gel of the double phosphorothioate modification of DNA-4 by isothermal and Whitfield extension showed similar extension results to the input of the single modification, dATP-αS or dTTP-αS, observed in **Figures 3.16** and **3.18**. All of the heat-cool cycle lengths showed the presence of the starting oligoseed below 75 bp which indicated that it had not all fully reacted during enzymatic extension. Furthermore, 5, 10 and 15 heat-cool cycles all extended to the same extent, with the maximum base pair length reached at 500 bp. The 20 heat-cool cycle extended product showed that much longer base pair lengths were achieved, with the maximum length at approximately 20,000 bp. The average length of this sample was visible at the brightest section of the size distribution at 500 bp. The longer extension length in this sample may be due to the production concatemeric formations. Further characterisation was collected by AFM imaging to determine whether the DNA had been extended to longer lengths or whether

other DNA structures were present. AFM of a 4 ng/μL sample of the 20 heat-cool cycle product with 0.5 mM $MgCl_2$ deposited onto freshly cleaved mica was imaged, **Figure 3.21**.

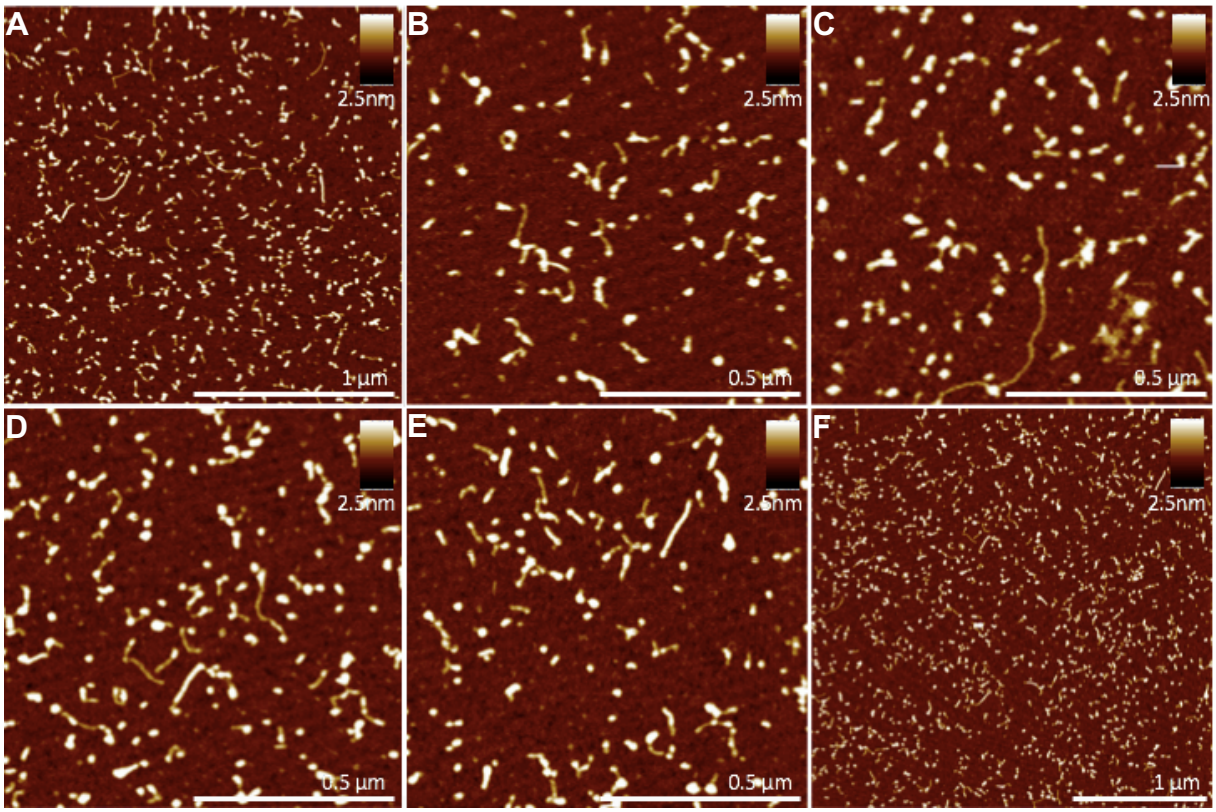


Figure 3.21. *AFM of Modified ABCD-DNA containing dATP-αS and dTTP-αS after 20 cycles enzymatic extension on freshly cleaved mica at a concentration of 4 ng/μL.*

The AFM images (**A**-**F**) confirmed that the double modification of the ABCD-DNA did not incorporate or extend as well as the single modifications. Some longer DNA strands are evident in the images, ~ 1,550 bp, however the majority of the DNA is very short < 100 bp. The average length of the ABCD-DNA generated with two modifications was 310 ± 240 bp (110 nm ± 80 nm). It was clear there was a height variation across the AFM mica surface as some parts of the DNA strands were much brighter than others, indicating higher heights on the surface as seen from the scale bar on the top right of the image. The average height was 1.2 ± 0.5 nm, however, the height of the DNA strands which appeared to be only dsDNA was 0.8 ± 0.2 nm and the height of the brighter areas was 1.7 nm ± 0.4 nm. This suggested that the brighter areas were a different structure to standard dsDNA as the height was higher, so the DNA in these areas consisted of aggregates, concatemers and self-annealed DNA.

To understand the difference between the two DNA species studied, computational calculations were performed by Joao de Souza (PhD Student, Bronowska Group, Newcastle University) comparing the phosphate backbone to the phosphorothioate backbone on the DNA helix. All computational modelling was calculated based on a 50 mer of the ABCD-DNA-4, with and without phosphorothioate modifications. The superficial charge calculated relates to the surface charge of the DNA backbone and would show if the presence of different atoms on the backbone would change the charge, **Figure 3.22**.

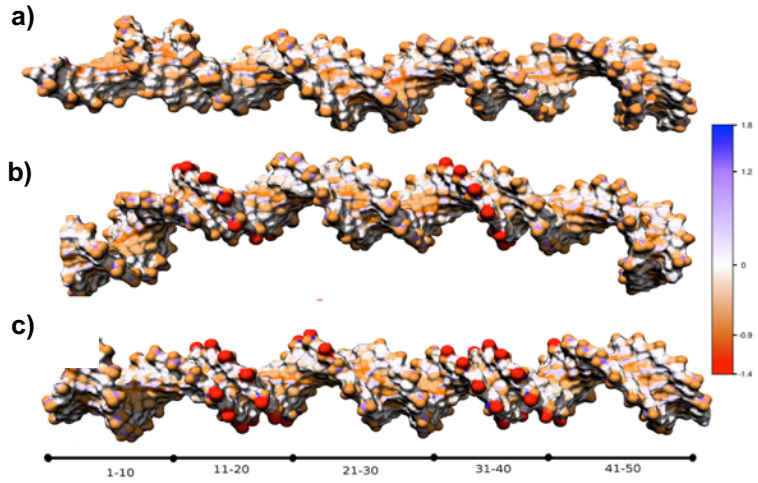


***Figure 3.22**. Superficial charges on ABCD-DNA. **a)** no modification **b)** dATP-αS modified **c)** dATP-αS and dTTP-αS modified.*

The scale bar on the right-hand side of **Figure 3.22** indicated the superficial charge, whereby the blue end of the scale was the most positive and the red end, the most negative. The replacement of oxygen (O) with sulfur (S), **Figure 3.22b and c**, showed that the charge on the backbone became more negative. This was indicated by the change in the outer most atom in the backbone shifting from orange on the DNA with no modification (**a**) to red on both of the modified backbones (**b and c**). This also meant that the dipole of the atom structure increased with the phosphorothioate incorporation and there was a bigger dipole difference between the phosphorous-sulfur (P-S) bond compared to the phosphorous-oxygen (P-O) bond. It was thought that these charges may have played a part in how the DNA was visualised on the AFM. **Figure 3.17** containing the dATP-αS modified DNA-4 showed DNA which was very aggregated/folded instead of individual vertically aligned DNA. In comparison, **Figure 3.14** which contained DNA-4 with no modifications, depicted DNA which was stretched out and although there was

some networking/branching it was easy to see the dsDNA strands with minimal folding of the DNA. The difference in the imaging between these two samples may have been caused by the difference in superficial charge as the sulfur would screen the Mg^{2+} ion better than the oxygen. Hence, the DNA would coil and fold more to be closer to the Mg^{2+} charge due to electronegative attraction. This screening effect did not only depend on charge but also the position of the atom on the backbone, so computational modelling of the atom position was studied, **Figure 3.23**.

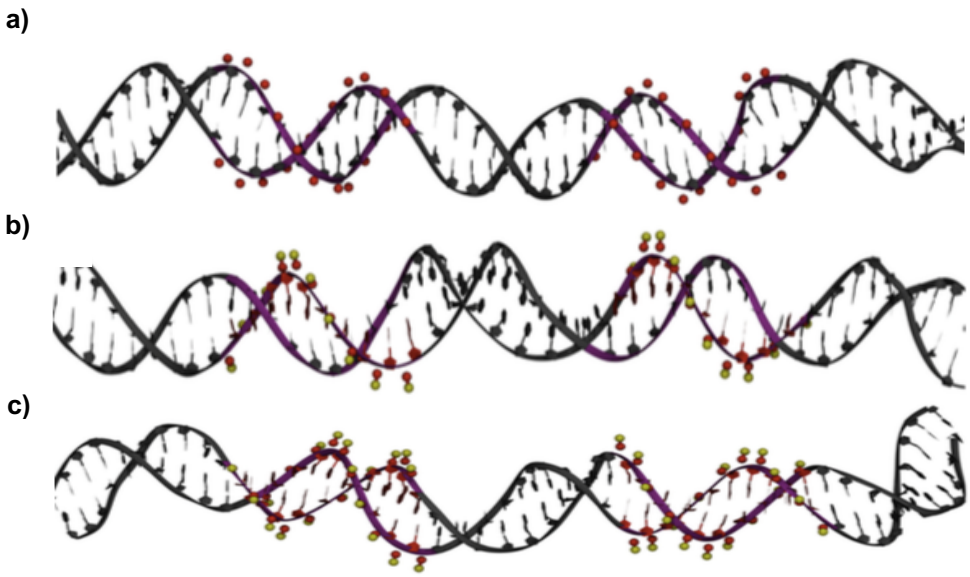


Figure 3.23. *Orientation of the modified backbone. Red= Phosphorous atoms. Yellow = Sulfur atoms.* ***a)*** *No modification* ***b)*** *dATP-αS modified* ***c)*** *dATP-αS and dTTP-αS modified.*

The computational modelling illustrated that the conformation of the helix backbone changed depending on the atom present, O or S. Compared with the O backbone, when S was present the helix shape distorted as observed in **Figure 3.23b and c**. It was also important to note that the S atoms tended to protrude more on the backbone than the O equivalents, hence the S atoms were more orientated towards the water than the O atoms. This made the phosphorothioate modified backbone a more attractive chemical handle for further functionalisation.

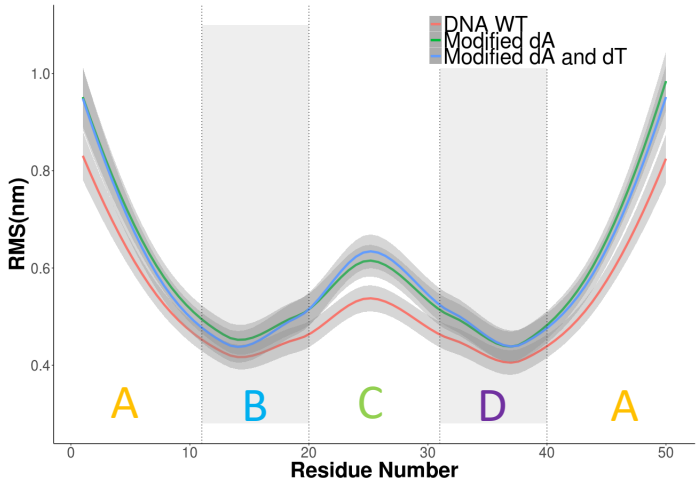
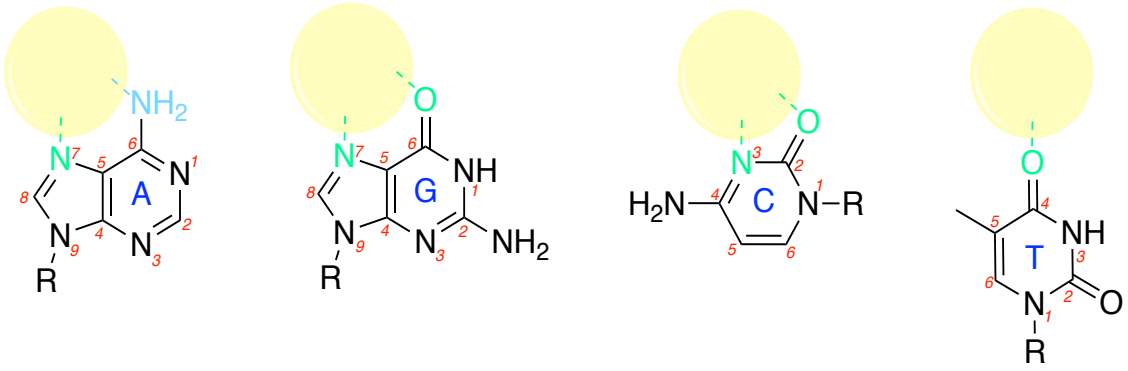


***Figure 3.24**. Spatial fluctuation per base for unmodified ABCD-DNA-4 and modification ABCD-DNA-4; 1 and 2 modifications.*

Spatial fluctuation of the helix was also computationally modelled with and without phosphorothioate modifications, both single and double, **Figure 3.24**. The shaded areas on the graph, B and D, showed the areas where dATP-αS and dTTP-αS modifications were incorporated for the modified backbone. It was clear that the backbone containing either single (green line) or double (blue line) phosphorothioate was calculated to be more flexible in its overall length compared to the standard backbone structure. C of the ABCD-DNA structure, P-O backbone, showed the highest increase in flexibility and may be the cause of the distorted helix structure discussed in **Figure 3.23**.

3.2.5 Modified DNA-4 AuNP Binding Studies

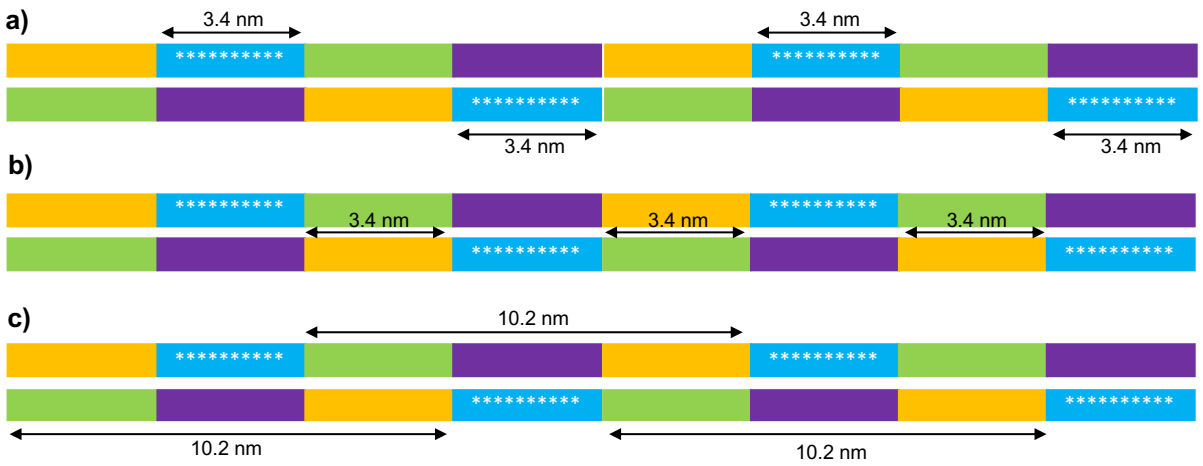
DNA binding to AuNPs has been manipulated for many years in applications such as DNA detection,[21] enhanced PCR specificity[22] and gold nanowire assembly.[23] A gold-thiol linkage is generally the favoured interaction due to the strong affinity of thiol for gold,[9] although it has also been proven that AuNPs can interact with the base,[24] nucleoside,[25] and the nucleotide.[26] Jang[27] reported that for dA, the exocyclic amino group (N6) and the nitrogen atom (N7) were responsible for the interaction between AuNPs and DNA. For dG, it was proposed that pyrimidine ring was the cause of the interaction, in particular (N6) and (O6). However, Pergolese *et al.*[28] suggested that it was due to (N7) of the imidazole ring. The interaction model for dC was through (N3) with some contribution of the ketonic oxygen (O2).[29] The dT base was reported as having the weakest interaction,[30] though some interaction with AuNPs was afforded from the ketonic oxygen (O4) which was closest to the methyl group.[29] All of these interactions are molecularly outlined in **Scheme 3.12**.



Scheme 3.12. *A, G, C and T base interactions with AuNP where the yellow sphere represents the AuNP and the dashed lines show the interaction. The green atoms illustrate strong interactions and the blue atoms weak interactions with the AuNP surface.*

Zhou *et al.*[9] reported stronger AuNP attachment than the standard bases through the incorporation of tandem phosphorothioate base modifications in the DNA sequence. It was thought that the presence of phosphorothioate DNA regions in a DNA sequence would promote preferential binding and increase the colloidal stability of the AuNPs.[9] Therefore, the phosphorothioate modified ABCD-DNA synthesised in **Section 3.2.4** is a good candidate for the generation of a DNA system that within its sequence design affords spatial control on AuNP binding.

The DNA-4 sequence was used for spatial binding studies as it offered good lengths of dATP-αS modified DNA after Whitfield extension (**Section 3.2.4**). This ensured that the distance between the AuNPs could be determined as there would be multiple AuNPs bound to the one dsDNA strand. The binding site was 10 bases (3.4 nm) in length and therefore, had space to bind one 3 nm AuNP. The spacing of the dATP-αS modified sections in the ABCD-DNA sequence was 30 bases on the single strand and 10 bases in duplex formation, **Scheme 3.13**. This meant there was a 3.4 nm gap between each binding site and so, there should be a measurable 3.4 nm gap between each AuNP if they have associated to the correct section of the DNA.



Scheme 3.13. *Spatial binding distances between modified base sites (indicated by *) on the ABCD-DNA-4 sequence.* ***a)*** *depicts the 3.4nm gap available for 3 nm AuNP binding at the modification site* ***b)*** *highlights the 3.4 nm spatial separation between modified sequences on the dsDNA* ***c)*** *shows the 10.2 nm spacing on the ssDNA between modification sites.*

The spatial binding of 3 nm gold nanoparticles (AuNPs), Nanopartz, USA, on dATP-αS modified DNA-4 (**Section 3.2.4**) was studied at various incubation times to determine the best spatial binding conditions. The spatial binding was visualised using AFM by the fabrication a standard AFM sample (20 heat-cool cycle sample of 4 ng/μL DNA with 0.5 mM $MgCl_2$ on freshly cleaved mica) and incubating the surface with 3 nm AuNPs for three time points; 15, 50 and 105 minutes. Unmodified extDNA-4 was used as the control sample. Initially, both samples were incubated with 3 nm AuNPs for 15 minutes before AFM images were collected, **Figure 3.25**.

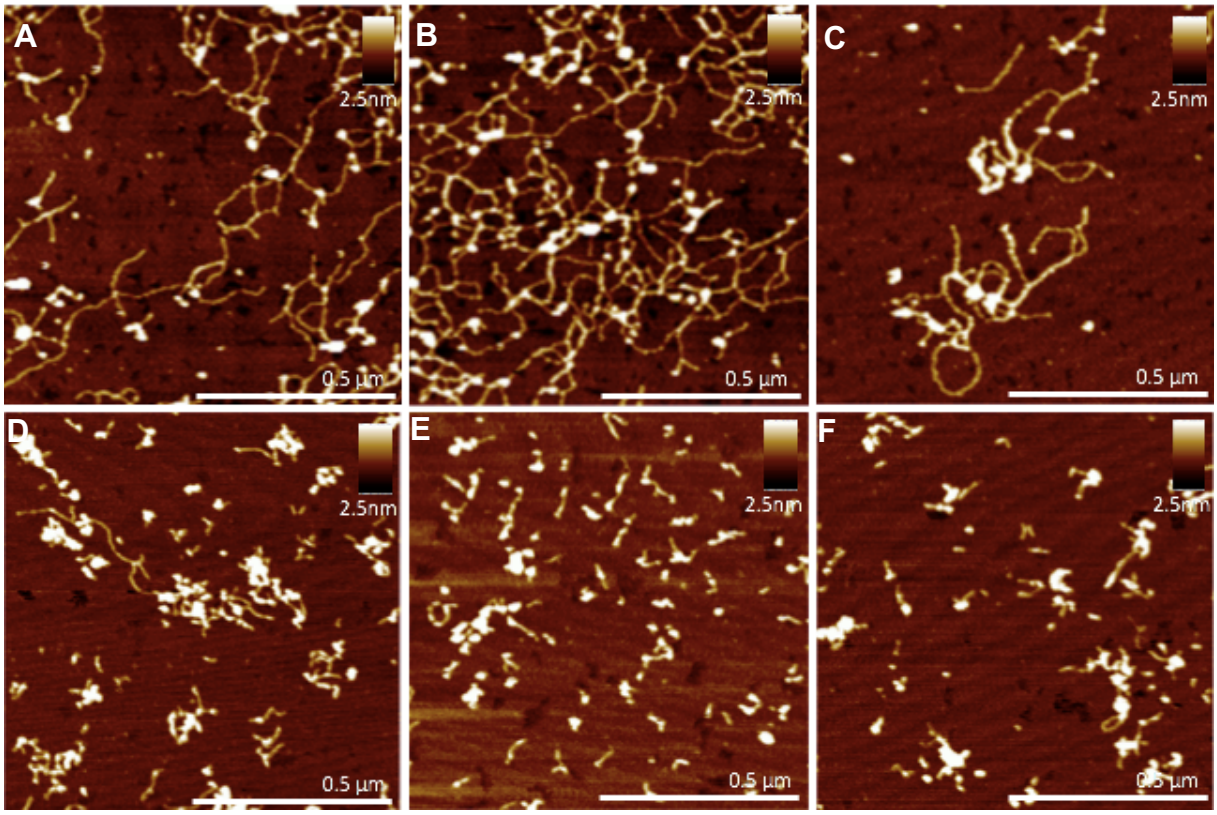


Figure 3.25. *AFM images after AuNP deposition for 15 minutes.* ***A-C:*** *unmodified ABCD-DNA-4* ***D-F:*** *dATP-αS modified ABCD-DNA-4.*

AFM images **A-C** depicted the unmodified DNA-4 after incubation with 3 nm AuNPs for 15 minutes. It was clear that there was some associated AuNPs to the DNA, however the DNA was mostly bare. It appeared that the AuNPs on this control sample, were only on small areas of the DNA, so binding only occurred where aggregation of the AuNPs was present. In comparison, the modified DNA-4 was slightly more covered with AuNPs, although there was still some bare DNA evident.

Due to the scan size of the AFM it was difficult to see whether the spatial binding was the correct length between particles, although from the AFM it was evident that the particles were associating more with the dATP-αS modified DNA. For the generation of a more covered DNA surface, the mica containing the DNA was soaked in the 3 nm AuNPs for 50 minutes before AFM imaging, **Figure 3.27**.

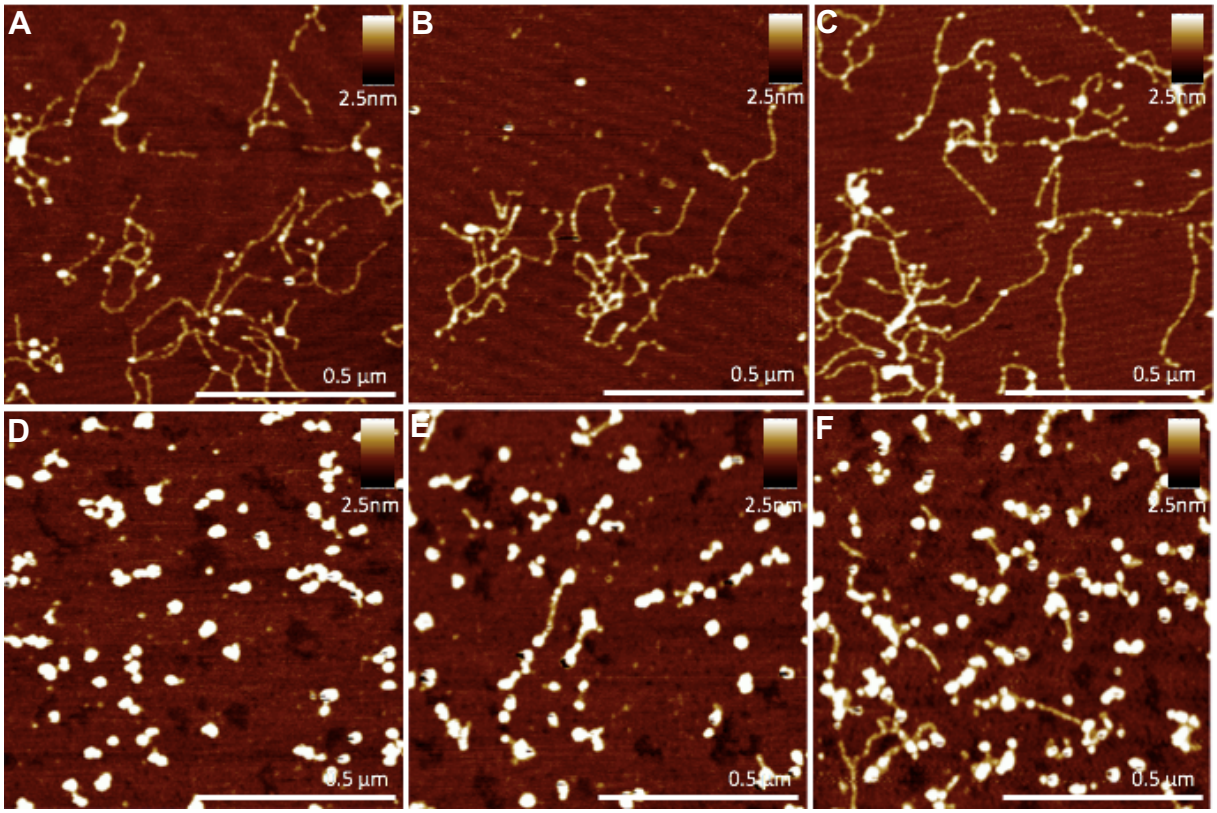


Figure 3.26. *AFM images after AuNP deposition for 50 minutes.* ***a)*** *unmodified ABCD-DNA-4* ***b)*** *dATP-αS modified ABCD-DNA-4.*

The difference in AuNP coordination between the unmodified ABCD-DNA-4 and dATP-αS modified ABCD-DNA-4, was easily observed after 50 minutes of 3 nm AuNP treatment. Images **A**-**C** showed the unmodified DNA-4 and mainly consisted of bare DNA with some AuNP aggregation present on some areas of the DNA. It was noticed that the structure of the unmodified DNA had changed after AuNP incubation and water washing as the DNA appeared to be continuous in length but visually more fragmented than observed in **Figure 3.14**, where no AuNP treatment had occurred. This fragmentation may be due to the DNA forming concatemeric structures during elongation to make long length DNA and the polymerase may not have filled the regions between concatemers to make continuous dsDNA.

In comparison, images **D**-**F**, which depicted the dATP-αS modified ABCD-DNA-4, showed DNA which was significantly more covered than the control sample, although there was still some bare DNA present. Furthermore, it was evident that the AuNPs had associated to the DNA in a more repetitive nature as would be expected from the spatial binding since it would be at a set distance of 3.4 nm.

Since the 3.4 nm spatial gap meant that a nanowire type structure should be observed on the AFM, **Figure 3.27** confirmed the reaction time for the AuNPs to associate with the DNA may need to be increased as it was clear that full AuNP coverage of the DNA was not achieved. Therefore, the DNA mica was reacted for 105 minutes with 3 nm AuNPs in an attempt to increase the DNA coverage, **Figure 3.30**.

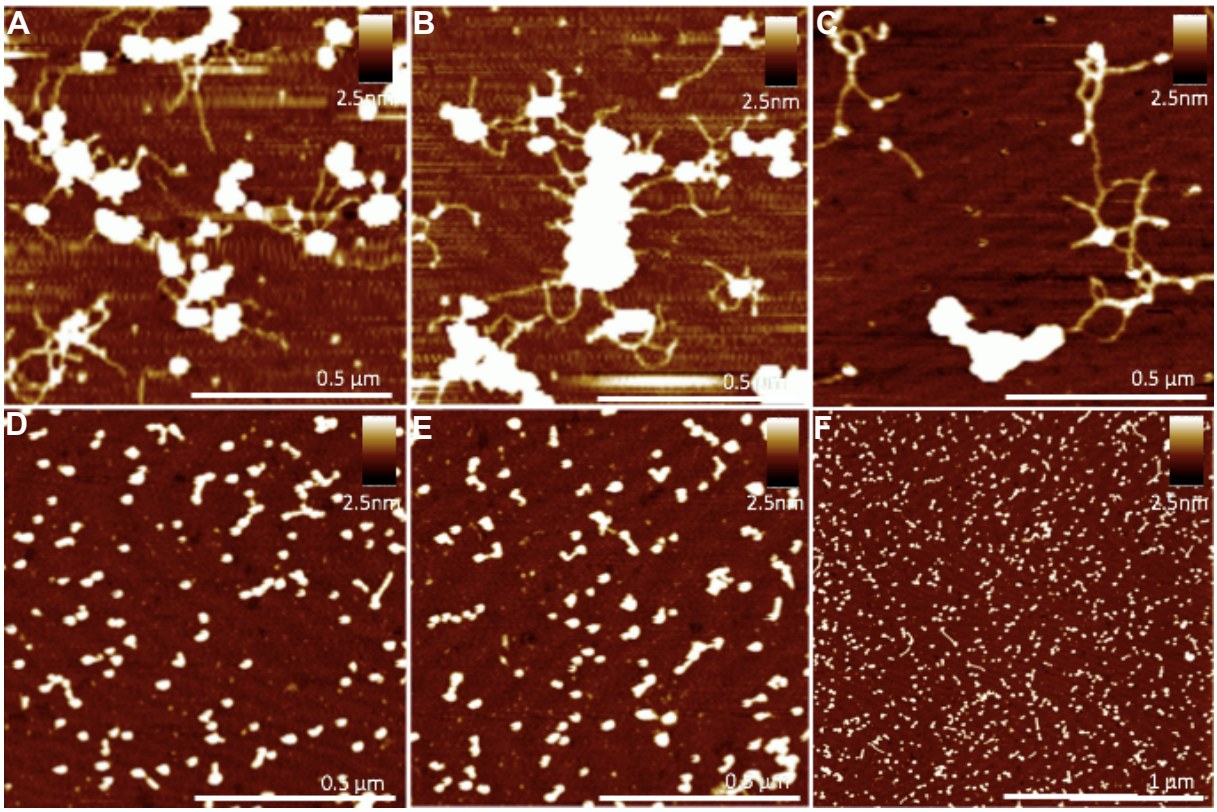


Figure 3.27. *AFM images after AuNP deposition for 105 minutes.* ***A-C:*** *unmodified ABCD-DNA-4* ***D-F:*** *dATP-αS modified ABCD-DNA-4.*

Images **A-C** contained the unmodified ABCD-DNA-4 with AuNP deposition after 105 minutes incubation and showed that the AuNPs had severely aggregated on the DNA backbone. It was clear that the 3 nm AuNPs only aggregated in areas where the DNA was present, so there must have been some association with the DNA backbone for the AuNPs to be localised to the DNA. There were still bare areas of DNA evident with an average height of 1.0 ± 0.3 nm, however the DNA-AuNP areas had an average height of 11 ± 1.5 nm, **Figure 3.31**. As Zhou *et al.*[9] suggested, this large increase in height from that expected (3 nm) was due to the aggregation of the AuNPs since the phosphate backbone reduced the colloidal stability of the AuNPs.

Images **A-C** were very different to those observed for the dATP-αS modified ABCD-DNA-4, (images **D**-**F**) as no aggregation of the AuNPs was observed after 105 minutes incubation. This suggested that the presence of the phosphorothioate on the DNA backbone reduced the aggregation effect after long incubation times. Although the dATP-αS modified ABCD-DNA-4 was much shorter in length than the unmodified ABCD-DNA-4, 390 bp compared to 710 bp respectively, the AuNPs were nicely aligned on the DNA and very little uncovered DNA was present in the sample.

In order to determine the optimum AuNP incubation time the 0, 15, 50 and 105 minute AFM images were compared for the unmodified ABCD-DNA-4 and the dATP-αS modified ABCD-DNA-4. The unmodified ABCD-DNA-4, **Figure 3.33**, showed that there was very little visual difference between 0, 15 and 50 minutes. The height of these three images were all similar and had an average height of 1.1 nm. Furthermore, the 105 minutes showed a difference visually in comparison to the shorter AuNP time incubations as it was obvious that the nanoparticles had severely aggregated. This was confirmed by the height of the nanoparticles in the area which were measured to be 11 nm, which was a difference of 8 nm compared to the expected size of 3 nm.

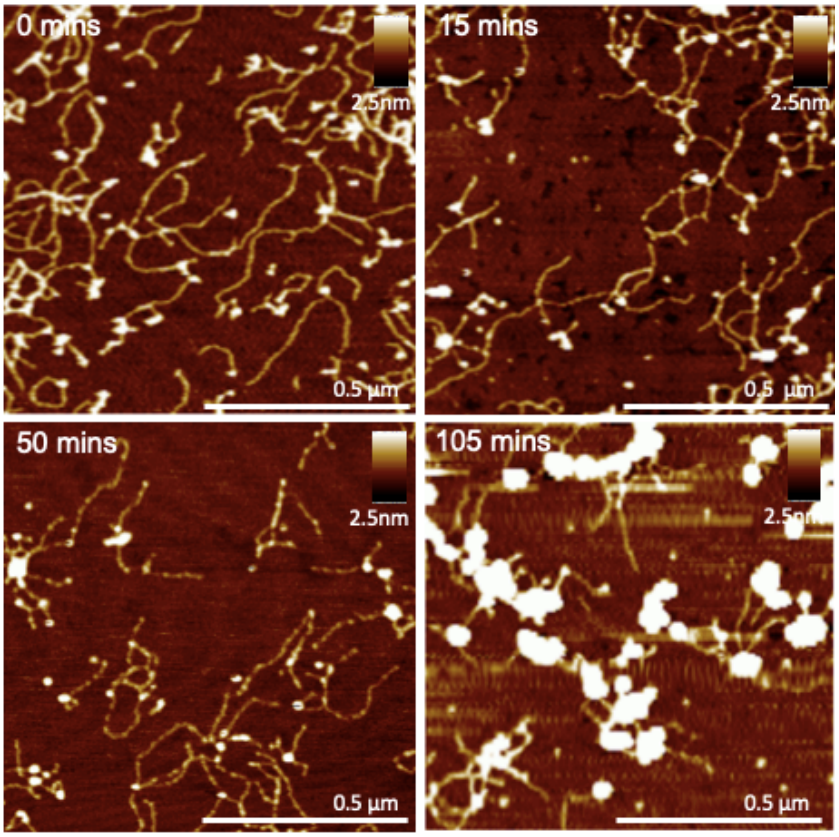


Figure 3.28. *ABCD-DNA comparison after AuNP deposition times of 0, 15, 50 and 105 minutes by AFM.*

Therefore, it was deemed that 105 minutes was too long an incubation time for the AuNPs. 15 or 50 minutes would give an ideal control result as they were both similar, however the optimum time would be dependent on the dATP-αS modified DNA backbone coverage with AuNPs, **Figure 3.34**.

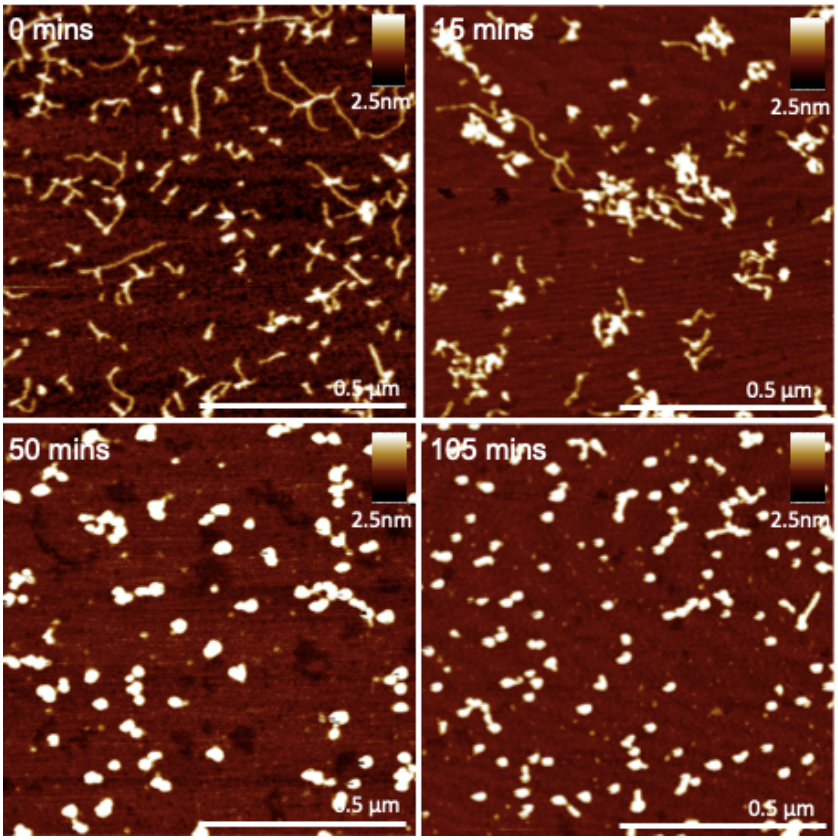


Figure 3.29. *ABCD-DNA-4 modified with dATP-αS comparison after AuNP deposition times of 0, 15, 50 and 105 minutes by AFM.*

It was evident here that a deposition time of 15 minutes did not offer much coverage of the DNA backbone with AuNPs and could not be used as the optimum incubation time. 50 and 105 minutes were visually very similar in coverage and height analysis as both had an average height of 3.3 nm. However, since the AuNPs aggregated very badly in the 105 minutes unmodified ABCD-DNA-4, it could not be used as a control sample. It was concluded that 50 minutes was the optimum time to bind 3 nm AuNPs to the phosphorothioate modified DNA backbone for spatial binding studies as it offered good DNA coverage, alongside a comparative control sample.

Although it was confirmed that binding of the AuNPs was successful and more preferential to the dATP-αS modified ABCD-DNA-4, the spatial binding could not be analysed by AFM as the scan size was too large. Tunnelling electron microscopy (TEM) was used to obtain detailed images of the single particles bound to the DNA. Since the DNA would not be seen by this technique, it was expected the AuNPs would be imaged as aligned in the shape of the DNA observed in the AFM images. The samples were prepared at 4 ng/µL of relevant DNA and incubated with 3 nm AuNPs for 50 minutes. The DNA-AuNP solution was then dropcast onto a holey carbon TEM grid and dried under UV

before imaged by Tracey Davey (Experimental Scientific Officer, Electron Microscopy Research Services, Newcastle University), **Figure 3.33**.

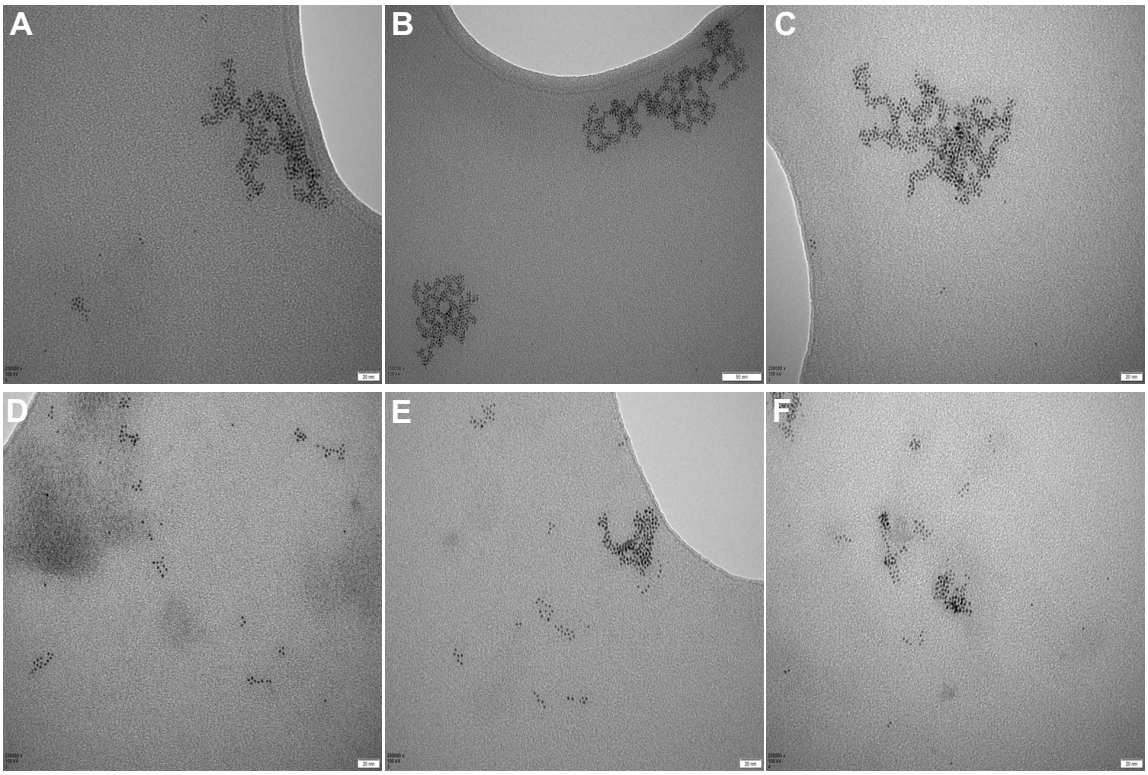


Figure 3.30 *TEM images after AuNP deposition for 50 minutes.* ***A-C****: unmodified ABCD-DNA-4* ***D-F:*** *dATP-αS modified ABCD-DNA-4.*

Images **A**-**C** exhibited unmodified DNA-4 with AuNP deposition and it was evident that there were large aggregations of particles present on the TEM grid surface. Controlled spacing of the AuNPs was not very apparent, although they were associated with the DNA. These images compared well to the observations found in the 50 minute AuNP deposition AFM images (**Figure 3.27, A**-**C**), where some association of the 3 nm AuNPs was apparent, however most of the DNA remained bare. Therefore, the aggregates observed in the TEM images (**A**-**C**) were linked to the brighter sections of DNA (average height 2.2 nm) measured in the AFM. TEM images **D**-**F** illustrated the dATP-αS modified DNA-4 with AuNP deposition and it was clear that some aggregation of the AuNPs had occurred, evident from the reduced spacing between the nanoparticles. This related to some areas of very dense DNA coverage observed in the AFM (**Figure 3.27, D**-**F**).

Furthermore, the AFM images indicated that not all of the DNA was covered in AuNPs and so a continuous alignment of AuNPs (similar to that of a nanowire) was not expected in the TEM images. There was alignment of the nanoparticles on the dATP-αS modified

DNA-4, most clearly visualised in image **D**, where it seemed to be AuNP association to the DNA followed by a section of bare DNA and then back to AuNP association. The spacing between the nanoparticles electrostatically bound to the DNA appeared to be more equally spaced than in images **A**-**C**. Full coverage of each modified base contained within the extDNA strand with 3 nm AuNPs was not achieved, however there was evidence in the TEM of linear structures compared to the unmodified DNA, although it could not be concluded from this data that intricate spatial binding was successful.

3.3 Conclusion

The aim of this chapter was to explore the versatility of the Whitfield enzymatic extension method. This was achieved by designing a new oligoseed compilation which was suitable for use in the Whitfield extension method but afforded a different structure of designer DNA to that previously described by Whitfield *et al.*[31] This new ABCD-DNA structure of DNA was designed for the eventual use in spatial binding studies as the design enabled the user to set the distance of their desired binding region within the initial DNA oligoseed before extension.

Successful enzymatic incorporation of a new sequence (D) by the design of an 3'-ABC-5'/5'-ABC-3' oligoseed containing three multi-base sequences was generated by using the Whitfield enzymatic extension method. For more difficult sequences, isothermal extension using Bst 2.0 polymerase was introduced before performing the Whitfield extension heat-cool cycles. The isothermal extension of the initial oligoseed increased the stability of the oligoseed by increasing both the length and Tm, making it more stable during the annealing step in the heat-cool cycle.

It was found that the ABCD-DNA method allowed for the incorporation of base modifications to introduce long regions (10 bp) of modified DNA at set spatial distances. Phosphorothioate modified bases (dATP-αS and dTTP-αS) were incorporated into the ABCD-DNA on both ssDNA and dsDNA. Although only one modification was discussed within this research, the ABCD-method offers flexibility to the user to change the sequence design without changing the initial oligoseed by selecting a different modification, either base or backbone modification. Therefore, it allows a library of modified sequences to be developed from the purchase of two oligos.

AuNP binding studies were performed to determine the use of specifically modified ABCD-DNA for spatial binding over standard ABCD-DNA. The AFM analysis confirmed that the 3 nm AuNPs specifically bound to the phosphorothioate regions of the DNA when compared to unmodified DNA. Unfortunately, spatial binding at the designed set distance was not achieved, as observed from the TEM imaging. However, the results from this chapter offered a positive outlook for the development of the ABCD-DNA method for spatial binding studies.

3.4 Experimental

3.4.1 Materials

Deoxynucleotide Preparation

Deoxynucleotides (**Table 3.9**) were purchased from Eurofins (Ebersberg, Germany).

Table 3.9. *Purchased deoxynucleotides*

Name	Origin	Sequence
DNA-1 A	Eurofins	5'-AAAAAAAAAGCCCCCCCCCCTTTTTTTTTC-3'
DNA-1 B	Eurofins	5'-CTTTTTTTTTCCCCCCCCCCGAAAAAAAAA-3'
DNA-2 A	Eurofins	5'-AAAATAAAATCCCCCCCCCCTTTTATTTTA-3'
DNA-2 B	Eurofins	5'-ATTTTATTTTCCCCCCCCCCTAAAATAAAA-3'
DNA-3 A	Eurofins	5'-AAAATAAAATCCCCCCCCCCCCCCCCCCCCTTTTATTTTA-3'
DNA-4 A	Eurofins	5'-CCCCCCCCCCAAAAAAAAAAGGGGGGGGGG-3'
DNA-4 B	Eurofins	5'-GGGGGGGGGGAAAAAAAAAACCCCCCCCCC-3'

Primer Template Annealing

The oligo seed duplexes (2μM) (**Table 3.10**) were prepared by the addition of DNA annealing buffer (10 mM HEPES pH 7.5, 100 mM sodium chloride (NaCl) and 1 mM ethylenediaminetetraacetic acid (EDTA)) to 10 μL of each deoxynucleotide required. The sample was heated to 95 °C for 10 minutes and cooled slowly to room temperature (25 °C). The annealed duplexes were then stored at -20 °C.

Table 3.10. *Summary of annealed duplexes.*

Name	Sequence
DNA-1	$[A_9GC_{10}T_9C]/[A_9GC_{10}T_9C]$
DNA-2	$[(A_4G)_2C_{10}(T_4C)_2]/[(A_4T)_2C_{10}(T_4C)_2]$
DNA-3	$[(A_4G)_2C_{20}(T_4C)_2]/[(A_4G)_2C_{10}(T_4C)_2]$
DNA-4	$[C_{10}A_{10}T_{10}]/[C_{10}A_{10}T_{10}]$

Table 3.11. *Description of DNA polymerases for enzymatic extension*

DNA Polymerase	Properties	Source
Thermococococcus gorgonarius polymerase B (Pfu-Pol) Z3 exo-	Archael family B Polymerase Low fidelity variant with the 5'→3' exonuclease activity removed and alterations to the fingers domain	Gifted from Dr. Colette Whitfield (PDR)
Pyrococcus polymerase B Deep Vent exo-	Family B polymerase with 5'→3' exonuclease activity removed	New England BioLabs Inc.

Table 3.12. *Description of DNA polymerases for isothermal extension*

DNA Polymerase	Properties	Source
Bst 2.0 Warm Start polymerase	Homologue of Bacillus stearothermophilus DNA polymerase I. Contains 5'→3' DNA polymerase activity with strong strand displacement and 5'→3' exonuclease activity removed.	New England BioLabs Inc.

3.4.2 Enzymatic Extension Method and Purification

0.2 µM DNA duplex, 200 nM polymerase, 10x reaction buffer (**Table 3.13**), 2 mM magnesium sulfate ($MgSO_4$) and 0.5 mM of each deoxynucleotide triphosphate (dNTPs dCTP, dGTP, dATP and dTTP) were mixed thoroughly and made up to the desired volume with nanopure water.

Thermocycling was performed using a VWR thermocycler by the method detailed below:

	30 seconds at 95 °C
N (number of cycles) x	30 seconds at 55 °C
	120 seconds at 72 °C

The products were cooled to 4 °C after the reaction and held at this temperature until removed from the thermocycler.

The extended DNA products were purified using either a QIAGEN QIAquick PCR purification kit (QIAGEN, Manchester) or a Monarch PCR and DNA clean up kit (NewEngland Bio Labs Inc., UK). Manufacturer's protocol was followed.

Table 3.13. *Polymerase Reaction Buffer*

Buffer	Components
10x Pfu-Pol Z3 exo -	200 mM Tris-HCl (pH 8.8 at 25 °C), 100 mM ammonium sulfate ($(NH_4)_2SO_4$), 100 mM potassium chloride (KCl), 1 % Triton X-100, 1 mg/mL bovine serum albumin (BSA) and 20 mM magnesium sulfate ($MgSO_4$)
1x ThermoPol	20 mM Tris-HCl (pH 8.8 at 25 °C), 10 mM ammonium sulfate ($(NH_4)_2SO_4$), 10 mM potassium chloride (KCl), 2 mM magnesium sulfate ($MgSO_4$) and 0.1 % Triton X-100

3.4.3 Isothermal Extension

0.2 µM DNA duplex, 320 u/mL Bst 2.0 Warm Start polymerase (New England Bio Labs Inc., UK), 10x isothermal reaction buffer, 6 mM magnesium sulfate ($MgSO_4$) and 1.4 mM of each deoxynucleotide triphosphate (dNTPs – dCTP, dATP, dTTP and dGTP) were mixed thoroughly and made up to the desired volume with nanopure water. Phosphorothioate modified bases (dATP-αS and dTTP-αS (Jena Bioscience, Germany) were used instead of the standard bases for modification of the DNA backbone.

Thermocycling was performed using a VWR thermocycler by the method detailed below:

Step 1	120 seconds at 95 °C
	60 seconds at 25 °C
Step 2	1 hour at 65 °C
Bst 2.0 enzyme added	120 seconds for 95 °C
	60 seconds for 25 °C
Step 3	1 hour at 65 °C
Bst 2.0 enzyme added	Hold at 25 °C

The isothermal products were used for enzymatic extension detailed in **Section 3.4.3**.

3.4.4 UV-Vis Spectroscopy

Standard

UV-Vis spectroscopy was carried out using a Thermo Scientific NanodropONE. 1.5 µL of sample was placed onto the stage and analysed between 1100 and 190 nm.

Melting Temperature

Melting curves were obtained on a Varian Cary100 Bio UV-visible spectrophotometer with a Varian Temperature controller. The 3 uM DNA solutions were pre-annealed in 0.1 M sodium phosphate buffer. The temperature increased from 5 °C to 95 °C at a rate of 1 °C/min. The absorbance was monitored at 260 nm and data collection was recorded at 0.5 °C intervals.

3.4.5 Agarose Gel Electrophoresis

Standard

DNA extension products were analysed by gel electrophoresis in TBE (tris, boric acid and EDTA) buffer. 1 % agarose (Melford Biolaboratories Ltd., Ipswich, UK) (1 g) was added to 1x TBE buffer (100 mL) and heated to dissolve. DNA samples were supplemented with 6x loading dye (2.5 % Ficoll-400, 11 mM EDTA, 3.3 mM Tris-HCl (pH 8 at 25 °C), 0.017 % sodium dodecyl sulfate (SDS) and 0.015 % bromophenol blue) (Thermo Scientific, UK) and GeneRuler 1 kilo base plus DNA ladder was used (Thermo Scientific, UK). The gels were run at 100 V, 100 mA, 10 W, for approximately 1 hour and post stained with 1 µg/µL ethidium bromide solution for 30 minutes. It was visualised using an ultra-violet transilluminator and imaged using UVIproMW software, version 11.02.

Lonza™ Flashgel Electrophoresis

Lonza™ gel electrophoresis was performed using 1.2 % agarose pre-cast gel kit and run as stated in the manufacturer's handbook. DNA was recovered from the recovery wells by addition of QIAGEN elution buffer (QIAGEN, Manchester, UK) to the wells. The recovered DNA fractions were purified using a QIAquick PCR purification kit (250) following the manufacturer's protocol

3.4.6 DNA Digestion

0.2 mg snake venom phosphodiesterase, 100 units bacterial alkaline phosphate (BAP) in 10 mM magnesium chloride (pH 7.0) and 10 mM magnesium chloride were added to the dsDNA sample (1-2 absorbance units at 260 nm) and made to a volume of 100 μL with nanopore water. The sample was incubated at 37 °C for 16-18 hours. 50 μL of digested product was analysed by reverse phase HPLC as per **Section 3.4.6**.

3.4.7 HPLC

HPLC was performed on a Waters 2487 machine using Picolog software, and a Varian Pro Star using a C_{18} reverse phase column for both systems. Buffer C was 0.1 M triethyl ammonium acetate at pH 6.5 with 2.5 % acetonitrile, and buffer B was 0.1 M triethyl ammonium acetate at pH 6.5 with 65 % acetonitrile at the gradient 0-30 % B over a period of 25 minutes. The gradient was returned to 0% B after 25 minutes) at a rate of 1ml/min. Absorbance was monitored at two wavelengths: 254 nm and 340 nm.

3.4.8 AFM

The top layer of the mica surface was cleaned using parcel tape. 2.5 μL of DNA sample was added to the 2.5 μL of 2mM magnesium chloride ($MgCl_2$) solution. The 5 μL mixture was spread on to the mica surface using a micropipette. After 5 minutes, the surface was washed with 5 μL of nanopure water and nitrogen gas was passed over the surface to remove excess water. The sample was dried for 30 minutes in the lamina flow fume cupboard. Light microscopy was used to locate the sample and cantilever position. AFM images were collected using a Dimension V with a Nanoscope controller (Veeco Instruments Inc., Metrology Group, Santa Barbara, CA). The tapping mode was used with an etched silicon tip (Tap 300 Al-G, 200 kHz, 40 N/m) on an isolation table. Nanoscope 7.00b19 software was used to acquire data.

3.4.9 Sequencing

Sequencing was performed by GATC biotech (Eurofins, Germany). 20-80 ng/μL of extended DNA after 5 cycles was prepared in 20 μL. 5 μL of 5 μM primer DNA (complementary DNA sequence) was prepared and both Eppendorf tubes sent for analysis.

3.4.10 Computational Simulations

First, the phosphorothioate modified backbone was created on the structure based on the canonical DNA oligo by swapping one of the oxygen atoms bounded to the phosphorous, to a sulfur atom. The phosphorothioate base charges were calculated by the AM1-BCC[32] method, using ACPYPE.[33] The configuration parameters (bonds, angles and dihedrals) were parametrised using AMBER99SB-ILDN[34] force field. The bond parameter between S-P and the angle parameters for S-P-O were extracted from GAFF.[35] The three final double helices were built in UCSF-Chimera ($C_{10}A_{10}G_{10}T_{10}G_{10}$ for no modification, $C_{10}A_{10}G_{10}T_{10}G_{10}$, where A was phosphorothioate modified for single strand and $C_{10}A_{10}G_{10}T_{10}G_{10}$, where A and T were phosphorothioate modified for double strand modification).

All molecular dynamic (MD) simulations were performed using GROMACS 5.1.[36] All of the different configurations were solvated with the TIP3P water model. Box distance was set to 1 nm and periodic boundary conditions were applied. Na^+ and Cl^- ions were added to achieve a 0.1 M concentration. The solvated systems were energy minimised and equilibrated. Minimisation ran using the steepest descent for 1,000 cycles followed by the conjugate gradient. Energy step size was set to 0.001 nm and the maximum number of steps was set to 50,000. The minimisation was stopped when the maximum force decreased below 1000 kJ/mol/nm using the Verlet cut-off scheme. Treatment of the long-range electrostatic interactions was set to Particle Mesh-Ewald (PME),[37] and the short range electrostatic and Van der Waals cut-off set to 1.0 nm. After the minimisation, NVT was performed for 20 ps with a time step of 2 fs and position restraints applied to the backbone. The constraint algorithm used was LINCS, which was applied to all bonds and angles in the system.[38] With the Verlet cutoff scheme and the non-bonded short range interaction, the cut-off was set to 1.0 nm. Long range electrostatics were again set to PME. The temperature coupling was set between the DNA and the non-DNA entities by using a Berendsen thermostat, with a time constant of 0.1 ps and the temperature set to reach 300 K with the pressure coupling off. NPT was run at 300 K with a Parrinello-Rahman pressure coupling on and set to 1 bar, without positional restraints. The MD simulations were performed for 50 ns per configuration.

3.4.11 AuNP Binding Studies

5 μL of 4 ng/μL extended DNA which contained 0.5 mM $MgCl_2$ was applied to a freshly cleaved mica surface and left for 5 minutes. The surface was washed with 5 μL of nanopure H_2O and dried with nitrogen. Samples were dried for a further 30 minutes on the lamina flow before imaged on the AFM.

After imaging the bare extended DNA, 5 μL of $9.2x10^{13}$ nps/mol 3 nm gold nanoparticles (Nanopartz, USA) were drop cast onto the surface and left for the amount of time desired. The mica surface was washed with 2 mL of nanopure water and dried with nitrogen. Samples were dried for a further 30 minutes on the lamina flow before imaged on the AFM.

3.4.12 TEM

5 μL of 4 ng/μL extended DNA was combined with 5 μL of $9.2x10^{13}$ nps/mol 3 nm gold nanoparticles (Nanopartz, USA) in a 0.5 mL Eppendorf and left for the amount of time desired.

2 μL of DNA:AuNP solution was dropcast onto a holey carbon TEM grid (EM Resolutions, Sheffield) and dried using a lamp as a heat source. The grids were examined on a Hitachi HT7800 transmission electron microscope using an Emsis Xarosa camera with radius software. Electron Microscopy Research Services, Newcastle University.

3.5 References

1. Whitfield, C. J. *et al.* Self-Priming Enzymatic Fabrication of Multiply Modified DNA. *Chem. - Eur. J.* **24**, 15267–15274 (2018).
2. Gan, R. *et al.* DNA phosphorothioate modifications influence the global transcriptional response and protect DNA from double-stranded breaks. *Sci. Rep.* **4**, 6642 (2015).
3. Wan, W. B. *et al.* Synthesis, biophysical properties and biological activity of second generation antisense oligonucleotides containing chiral phosphorothioate linkages. *Nucleic Acids Res.* **42**, 13456–13468 (2014).
4. Eckstein, F. Phosphorothioate Oligodeoxynucleotides: What Is Their Origin and What Is Unique About Them? *Antisense Nucleic Acid Drug Dev.* **10**, 117–121 (2000).
5. Sandström, P., Boncheva, M. & Åkerman, B. Nonspecific and Thiol-Specific Binding of DNA to Gold Nanoparticles. *Langmuir* **19**, 7537–7543 (2003).
6. Nicewarner Peña, S. R., Raina, S., Goodrich, G. P., Fedoroff, N. V. & Keating, C. D. Hybridization and Enzymatic Extension of Au Nanoparticle-Bound Oligonucleotides. *J. Am. Chem. Soc.* **124**, 7314–7323 (2002).
7. Zikich, D., Borovok, N., Molotsky, T. & Kotlyar, A. Synthesis and AFM Characterization of Poly(dG)-poly(dC)-gold Nanoparticle Conjugates. *Bioconjug. Chem.* **21**, 544–547 (2010).
8. Eidelshtein, G., Fattal, M., Avishai, G., Giannini, C. & Kotlyar, A. Preparation, Characterization and Manipulation of Conjugates between Gold Nanoparticles and DNA. *Nanomaterials* **6**, 167–175 (2016).
9. Zhou, W., Wang, F., Ding, J. & Liu, J. Tandem Phosphorothioate Modifications for DNA Adsorption Strength and Polarity Control on Gold Nanoparticles. *ACS Appl. Mater. Interfaces* **6**, 14795–14800 (2014).
10. Jiang, L. *et al.* Sterically mediated two-dimensional architectures in aggregates of Au nanoparticles directed by phosphorothioate oligonucleotide-DNA. *Adv. Mater.* **17**, 2066–2070 (2005).
11. Ghabboun, J. *et al.* Specific and efficient adsorption of phosphorothioated DNA on Au-based surfaces and electrodes. *Appl. Phys. Lett.* **91**, (2007).
12. Campos, R., Kotlyar, A. & Ferapontova, E. E. DNA-Mediated Electron Transfer in DNA Duplexes Tethered to Gold Electrodes via Phosphorothioated dA Tags. *Langmuir* **30**, 11853–11857 (2014).
13. Kumar, A. *et al.* Assembling gold nanoparticles in solution using phosphorothioate DNA as structural interconnects. *Curr. Sci.* **84**, 71–74 (2003).

14. Lubitz, I. & Kotlyar, A. G4-DNA-Coated Gold Nanoparticles: Synthesis and Assembly. *Bioconjug. Chem.* **22**, 2043–2047 (2011).
15. Borovok, N., Gillon, E. & Kotlyar, A. Synthesis and Assembly of Conjugates Bearing Specific Numbers of DNA Strands per Gold Nanoparticle. *Bioconjug. Chem.* **23**, 916–922 (2012).
16. Lee, J. H. *et al.* Site-Specific Control of Distances between Gold Nanoparticles Using Phosphorothioate Anchors on DNA and a Short Bifunctional Molecular Fastener. *Angew. Chem. Int. Ed.* **46**, 9006–9010 (2007).
17. Ghabboun, J. *et al.* Specific and efficient adsorption of phosphorothioated DNA on Au-based surfaces and electrodes. *Appl. Phys. Lett.* **91**, (2007).
18. Whitfield, C. J. Enzymatic Protocols for the Synthesis of Designer DNA. (Newcastle University, 2016).
19. Freifelder, D. & Davison, P. F. Hyperchromicity and Strand Separation in Bacterial DNA. *Biophys. J.* **2**, 249–256 (1962).
20. Hansma, H. G., Revenko, I., Kim, K. & Laney, D. E. Atomic force microscopy of long and short double-stranded, single-stranded and triple-stranded nucleic acids. *Nucleic Acids Res.* **24**, 713–20 (1996).
21. Goodman, C. M. *et al.* DNA-binding by Functionalized Gold Nanoparticles: Mechanism and Structural Requirements. *Chem. Biol. Drug Des.* **67**, 297–304 (2006).
22. Li, H. *et al.* Nanoparticle PCR: Nanogold-Assisted PCR with Enhanced Specificity. *Angew. Chem. Int. Ed.* **44**, 5100–5103 (2005).
23. Ongaro, A. *et al.* DNA-Templated Assembly of Conducting Gold Nanowires between Gold Electrodes on a Silicon Oxide Substrate. *Chem. Mater.* **17**, 1959–1964 (2005).
24. Gourishankar, A., Shukla, S., Ganesh, K. N. & Sastry, M. Isothermal Titration Calorimetry Studies on the Binding of DNA Bases and PNA Base Monomers to Gold Nanoparticles. *J. Am. Chem. Soc.* **126**, 13186–13187 (2004).
25. Storhoff, J. J., Elghanian, R., Mirkin, C. A. & Letsinger, R. L. Sequence-Dependent Stability of DNA-Modified Gold Nanoparticles. *Langmuir* **18**, 6666–6670 (2002).
26. Yang, J., Pong, B.-K., Lee, J. Y. & Too, H.-P. Dissociation of double-stranded DNA by small metal nanoparticles. *J. Inorg. Biochem.* **101**, 824–830 (2007).
27. Jang, N. H. The Coordination Chemistry of DNA Nucleosides on Gold Nanoparticles as a Probe by SERS. *Bull. Korean Chem. Soc.* **23**, 1790–1800 (2002).

28. Pergolese, B., Bonifacio, A. & Bigotto, A. SERS studies of the adsorption of guanine derivatives on gold colloidal nanoparticles. *Phys. Chem. Chem. Phys.* **7**, 3610–3613 (2005).
29. Jang, N. H. *The Coordination Chemistry of DNA Nucleosides on Gold Nanoparticles as a Probe by SERS*. *Bull. Korean Chem. Soc* **23**, (1790).
30. Carnerero, J. M., Jimenez-Ruiz, A., Castillo, P. M. & Prado-Gotor, R. Covalent and Non-Covalent DNA-Gold-Nanoparticle Interactions: New Avenues of Research. *ChemPhysChem* **18**, 17–33 (2017).
31. Whitfield, C. J., Turley, A. T., Tuite, E. M., Connolly, B. A. & Pike, A. R. Enzymatic Method for the Synthesis of Long DNA Sequences with Multiple Repeat Units. *Angew. Chem. Int. Ed.* **54**, 8971–8974 (2015).
32. Jakalian, A., Jack, D. B. & Bayly, C. I. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *J. Comput. Chem.* **23**, 1623–1641 (2002).
33. Sousa da Silva, A. W. & Vranken, W. F. ACPYPE - AnteChamber PYthon Parser interfacE. *BMC Res. Notes* **5**, 367–367 (2012).
34. Lindorff-Larsen, K. *et al.* Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins Struct. Funct. Bioinforma.* **78**, 1950–1958 (2010).
35. Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A. & Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **25**, 1157–1174 (2004).
36. Berendsen, H. J. C., Spoel, D. van der & Drunen, R. van. GROMACS: A message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun.* **91**, 43–56 (1995).
37. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092 (1993).
38. Hess, B. P-LINCS: A Parallel Linear Constraint Solver for Molecular Simulation. *J. Chem. Theory Comput.* **4**, 116–122 (2008).

Chapter 4

DNA Brush Surfaces for Sensing Applications

Contents

4.1 Introduction

Diseases such as cystic fibrosis[1] and colorectal cancer[2] originate from gene mutations causing a modification in the DNA sequence. The early detection of the gene defect and therefore, the disease, would allow for quicker patient treatment and increased quality of life. Furthermore, the fabrication of long DNA brush surfaces from DNA repeat sequences, has the potential to differentiate between individual sequences and detect single base mismatches in diseases. These qualities are ideal for the production of a bio-sensor device.

This chapter utilises siloxane surface chemistry alongside the Whitfield enzymatic extension, to produce long repeating DNA brush surfaces for sensing applications. More specifically, this involves the production of surfaces that target DNA sequences bearing single or multiple mismatches which are known mutations in colorectal cancer and cystic fibrosis. Initially, a sequence comprised of GATC repeats was studied to demonstrate the concept of an enhanced fluorescent signal by multiplexing the immobilised probe sequence in the z-direction. Secondly, the colorectal cancer MMR sequence, BAT25, where single base mutations can occur in a section of repeating bases in the sequence, was used to test the sensitivity and selectivity of the proposed approach. Finally, in continuation of previous research on the cystic fibrosis sequence,[3] the precision of the extended DNA length was investigated to help quantify the fluorescence response.

Fundamentally, the advantage of having multiple repeats of the probe sequence within one strand of the DNA is the increasing number of base pair binding sites available for the fluorescent intercalator, PG. This results in a higher density of fluorescent tags per unit area of the surface and therefore, improves the sensitivity and selectivity of the target sequences. The majority of DNA surfaces are generated by the addition of short DNA oligos, ~ 100 bp, with a number of recent examples adding DNA up to 1,000 bp. For example, Bracha *et al*.[4] generated a DNA density gradient across the surface by immobilising DNA, 968 bp in length, **Figure 4.1a**. The ionic strength was varied and the effect on the DNA length at the different densities across the surface was studied. Four DNA brush regimes were observed; osmotic, salted, mushroom and neutral, **Figure 4.1b**.

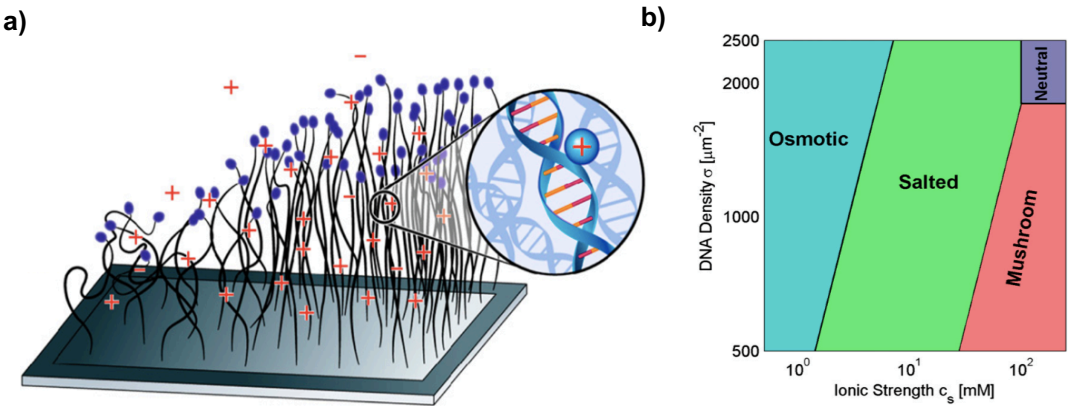
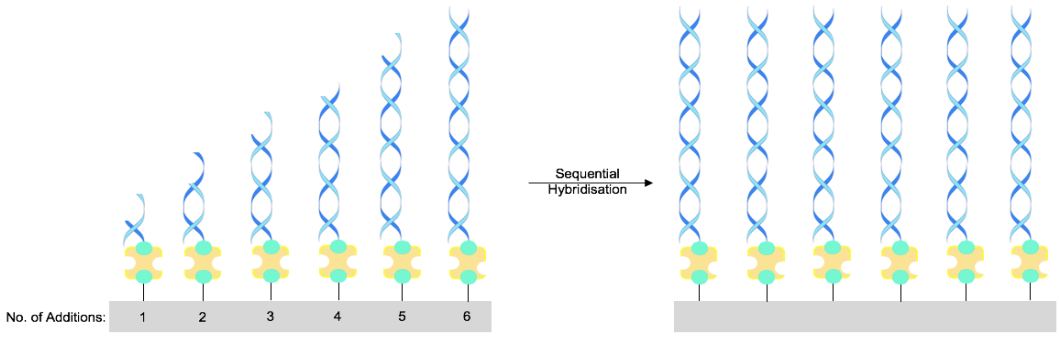


Figure 4.1. a) DNA brush density gradient across the surface **b)** *DNA brush regimes* produced with *differing DNA densities and ionic strengths.*[4]

The osmotic brush regime corresponded to forced elongation of the DNA strand, resulting from increased entropy of the counter ions within the brush structure. Reduction in the DNA height was observed in the salted regime when there was an imbalance in the ionic strength and therefore, entropy loss. The mushroom brush regime forms when the flexibility and folding capabilities of the DNA brush increased due to reduced chain interactions from high salt concentrations and low DNA packing density. Finally, the neutral regime was detected when there was high salt and high DNA density.[4] The Bracha group have evidently demonstrated the possibility of producing a DNA surface with DNA lengths > 100 bp, however the long DNA must be immobilised after synthesis which could therefore affect the surface packing density.

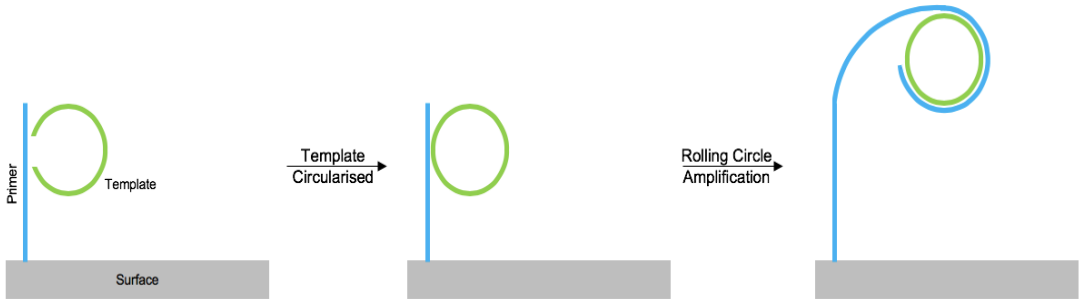
It would be advantageous to have the ability to grow short DNA oligos immobilised on the substrate in the z-direction away from the surface, thus increasing the packing density as the packing density would not be hindered on immobilisation of long DNA lengths due the short nature of the oligo probe strand. DNA growth after surface immobilisation was explored by the Åkerman group through sequential concatemer formation, **Scheme 4.1**.[5]



Scheme 4.1. *DNA brush surfaces fabricated by sequential oligo hybridisation to form long DNA lengths made up of concatemers.*[5]

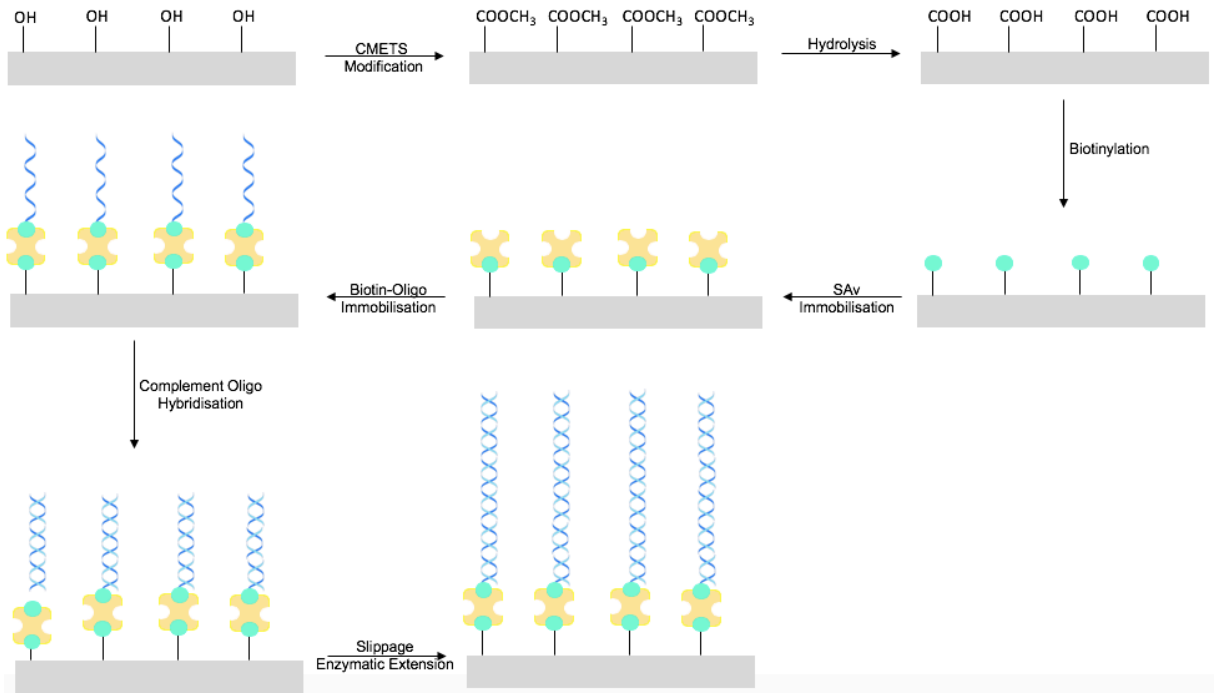
The immobilisation of a 59 bp biotinylated probe strand using a SAv modified surface and hybridisation of a shorter target complementary strand (34 bp) generated an overhang at the strand terminus. Each concatemer addition step added approximately 25 bp to the DNA layer, so sequential hybridisation of the oligos was used to build up concatemeric DNA layers of 534 bp in length. For the purpose of DNA sensing applications, this method is limited as the concatemeric surface would not be able to be dehybridised to allow for target probe hybridisation as the brush structure would dismantle.

RCA is an alternative method for the extension of short immobilised DNA probe strands off a solid surface.[6,7] The ssDNA probe strand is anchored to a solid support and the complement ssDNA template is hybridised to the immobilised probe. The template is circularised by DNA ligase and the primer is extended by the use of a polymerase, **Scheme 4.2**. The RCA greatly extended the DNA primer length, with reported values of > 726 bp from Hatch *et al.*[6]



Scheme 4.2. *RCA of the ssDNA probe immobilised on the surface.*[6,7]

The final section of this chapter focusses on the fabrication of photo-induced patterned DNA brush surfaces, in collaboration with the Ijiro group (Hokkaido University, Sapporo). DNA brushes contain a layer of DNA polymer chains which are densely attached onto a substrate and hence exhibit many novel properties such as wettability, adsorption and lubrication.[8–10] The Ijiro group utilised the SAv-biotin interaction to attach DNA to the substrate surface and extend the DNA length by surface-initiated enzymatic slippage extension using Klenow fragment, **Scheme 4.3**.[11,12] The extended DNA fragments had lengths of 9,700 bp after 60 minutes extension at 37 °C.[11]



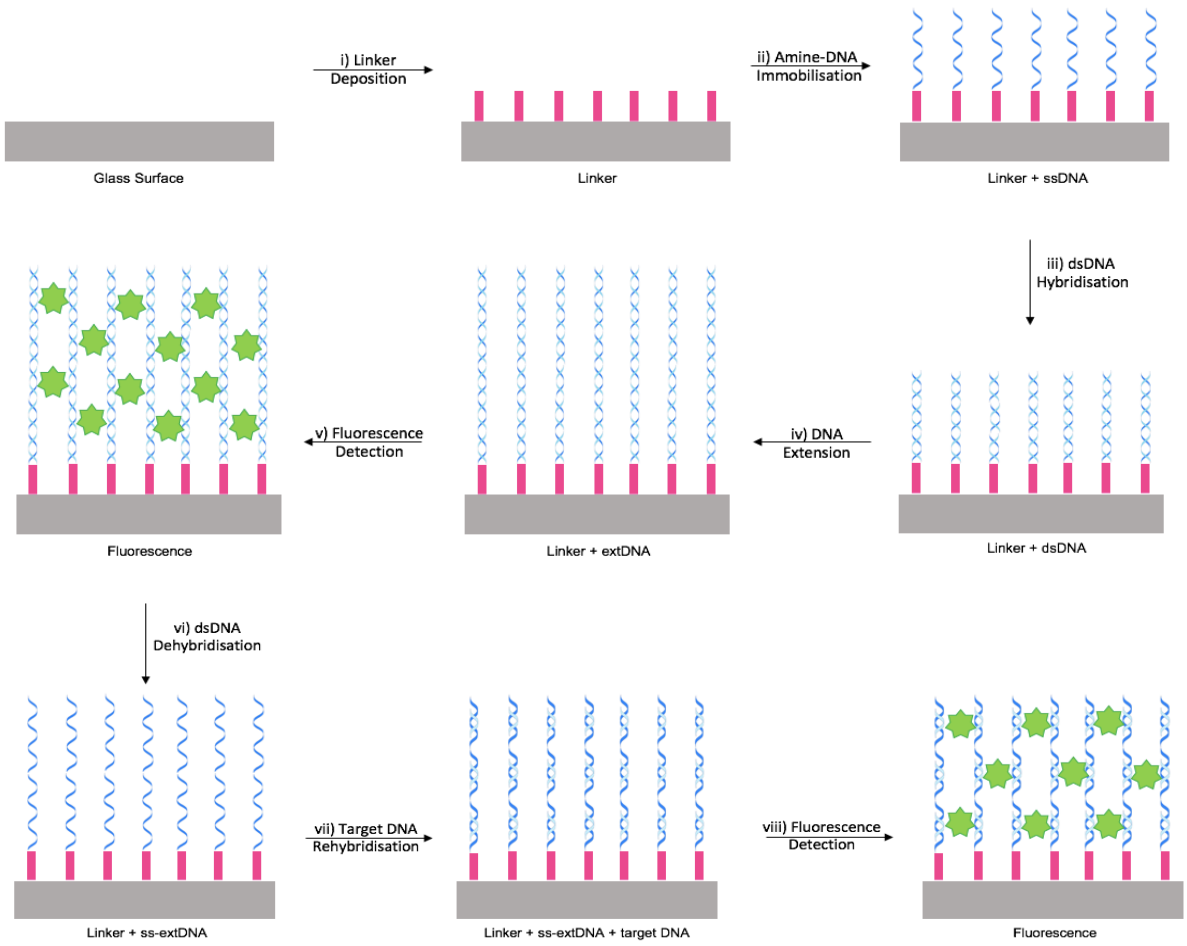
Scheme 4.3. *Streptavidin-biotin surface immobilisation, DNA attachment and slippage extension.*[11,12]

The main application for DNA brush surfaces is to create a sensing platform. The introduction of the brush structure would increase the target DNA loading and therefore enhance the output signal by the increase in the number of fluorescent intercalator binding sites. This device would remove the need for expensive fluorescent labelling of the target DNA oligos and produce a simple, cost-effective diagnostic device.

4.2 Results and Discussion

The fabrication of an extended DNA surface is illustrated in **Scheme 4.4**. This protocol builds on previous DNA immobilisation work described in **Section 4.1**. In summary, the glass surface is modified with an aldehyde 3' terminated linker (APEGDMES) and the ssDNA probe strand is immobilised. The ssDNA target strand is hybridised to the probe strand to form an immobilised oligo seed suitable for undergoing the Whitfield enzymatic extension. After enzymatic extension on the surface, the non-immobilised long ssDNA complement strand is dehybridised and the short target DNA rehybridised. PG dye intercalates with the dsDNA and the fluorescence of the surface is measured by fluorescence microscopy to determine the presence of the target DNA. This dye was chosen due its ease of excitation at 496 nm using fluorescence microscopy.

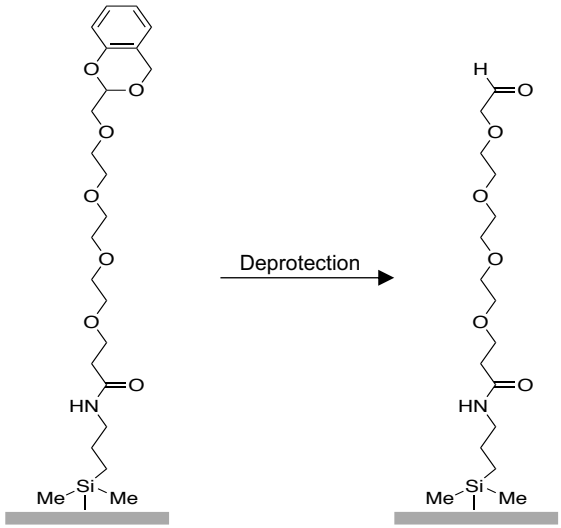
The reproducibility of the dehybridisation/rehybridisation process has proven difficult to obtain in previous research[3] due to the inability to guarantee replication of the DNA chips and the hybridisation process. This is caused by a large number of variables involved in the surface fabrication as each chip will vary in packing density and DNA length. Furthermore, the structure of the DNA on the surface may vary between DNA chips as the salt concentration and localised pH will affect the DNA.[4] However, optimisation of the hybridisation and rehybridisation conditions could improve the confidence level in the protocol to minimise the influence of the main sources of error. It is important for these conditions to be reproducible as the target DNA detection determines the sequence of the hybridised DNA; complementary, non-complementary, single, double or triple base mutation.



Scheme 4.4. *Surface fabrication and Whitfield enzymatic extension to enable rehybridisation of the target DNA for disease detection. i) APEGDMES linker deposition ii) AmC6-ssDNA immobilisation with NaCNBH3 iii) dsDNA hybridisation iv) Whitfield enzymatic extension v) Fluorescence measured vi) Surface dehybridised vii) Target DNA rehybridisation of the surface viii) Fluorescence measured.*

4.2.1 Surface Extension – $[GATC]_5$/$[GATC]_5$

The probe DNA sequence (5'-AmC_6-$[GATC]_5$-3') and the $[GATC]_5$ target sequence were used as a proof of concept oligoseed to ensure extension off the surface was successful, before more complicated real-life sequences were introduced. According to the method developed by Lunn,[3] DNA modified surfaces were fabricated by using an acetalpolyethyleneglycoldimethylethoxysilane (APEGDMES) linker, **Scheme 4.5**.

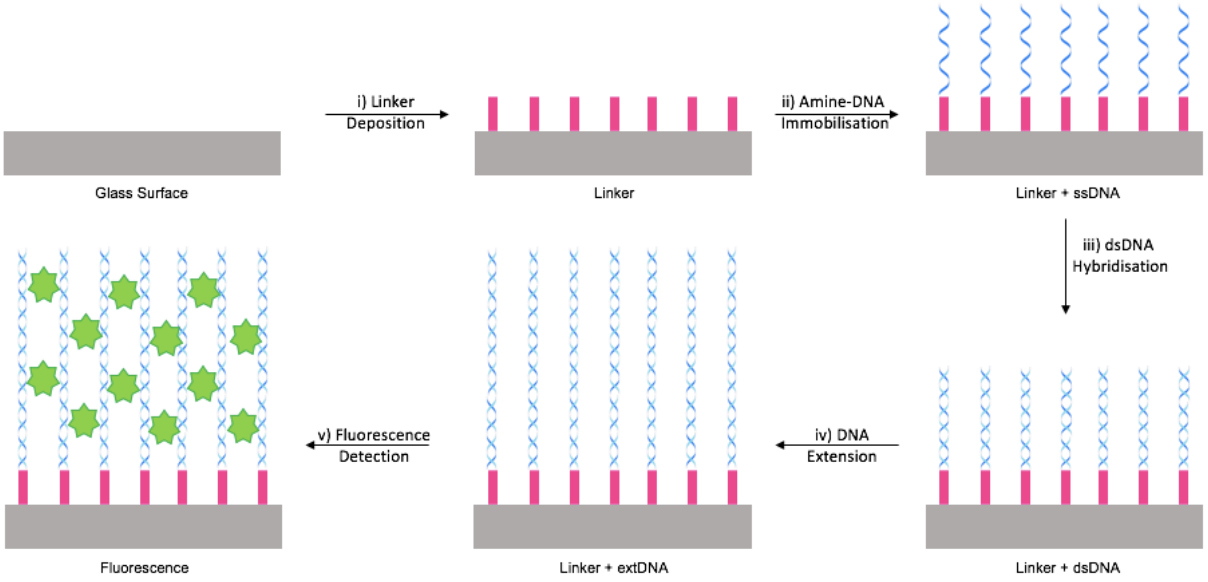


Scheme 4.5. *APEGDMES linker molecule.*

Four polyethylene glycol (PEG) groups were integrated within the APEGDMES linker molecule to decrease biofouling and thus avoid non-specific DNA adsorption onto the glass surface.[13,14] Furthermore, this PEG_4 linkage increased the distance of the terminal aldehyde from the surface which would ensure there was no steric hindrance on covalent DNA attachment. The APEGDMES linker was dissolved in toluene and surfaces were immersed in the linker overnight at 65 °C. This allowed for the free hydroxyl groups on the surface to react with the siloxane end of the linker (silanisation reaction).[15] This linker provided facile surface fabrication, as only one reaction step was required prior to DNA immobilisation.

Full fabrication of the surface is outlined in **Scheme 4.6**. The surface substrate was generated by cutting glass slides into 0.5 cm^2 squares and cleaning by sonication in acetone, isopropanol (IPA) and nanopure water for 15 minutes, followed by 10 minutes O_2 plasma treatment. This step was employed to remove any residual organics on the surface and activated the surface hydroxyl groups for linker modification. The APEGDMES linker in toluene was reacted with the newly formed oxide layer for 16 hours at 65 °C. After washing, the linker modified glass surface was treated with a 10 % acetic acid solution to deprotect the aldehyde which contained the amino-modified DNA oligo (probe strand). Reductive amination between the new terminal aldehyde functionality and the amino-modified DNA was initiated by the addition of sodium cyanoborohydride ($NaBH_3CN$). The short ssDNA probe strand was covalently attached to the linker molecule and a 30 minute 0.5x PBS buffer wash of the surface ensured the removal of any physisorbed amino-modified DNA. The immobilised ssDNA surface was hybridised with

the complementary ssDNA oligo (target strand), which produced a 20 bp dsDNA oligoseed on the surface. The presence of the oligoseed on the surface allows for the DNA to be elongated in the z-direction (away from the surface) by the Whitfield enzymatic extension method, as described in **Chapter 1**, to form DNA brush surfaces.



***Scheme 4.6**. DNA surface fabrication using APEGDMES linker and amine terminated DNA. i) APEGDMES linker deposition ii) AmC_6-ssDNA immobilisation with $NaCNBH_3$ iii) dsDNA hybridisation iv) Whitfield enzymatic extension from the surface v) Fluorescence detection of hybridised DNA surface.*

The dsDNA surface was enzymatically extended using *Thermococcus gorgonarius* polymerase B (Pfu-Pol) Z3 exo$^-$ for a chosen number of heat-cool cycles during the Whitfield enzymatic extension. The generation of the extDNA brush surface was observed by the addition of a fluorescent intercalator, PG. As discussed in **Chapter 2**, PG displays an increase in fluorescence emission upon binding to dsDNA[16] and therefore, allowed an increase in fluorescence intensity to be observed after enzymatic extension of the dsDNA due to the increased number of binding sites.

After each modification step, a 200-fold dilution of PG solution was dropcast onto the surface and left for 20 minutes before it was washed with nanopure water and dried with N_2. The surface was analysed by fluorescence imaging, either by fluorescence microscopy or on a molecular imager. Due to the absorption of PG by the glass surface itself, it was difficult to determine the fluorescence visually by eye, so the collected fluorescent images were processed using Image J to calculate the integrated density of the surface. The integrated density is defined as the sum of pixel values in the image and is the equivalent to the product of area and mean grey value.

The [GATC] immobilised surfaces were extended for 20 heat-cool cycles since these conditions provided good extDNA lengths in solution, ~ 1,000 bp.[17] The extDNA surfaces were imaged by fluorescence and analysed by Image J. The average integrated density for each modification step was represented in a bar chart for ease of comparison, **Figure 4.2**.

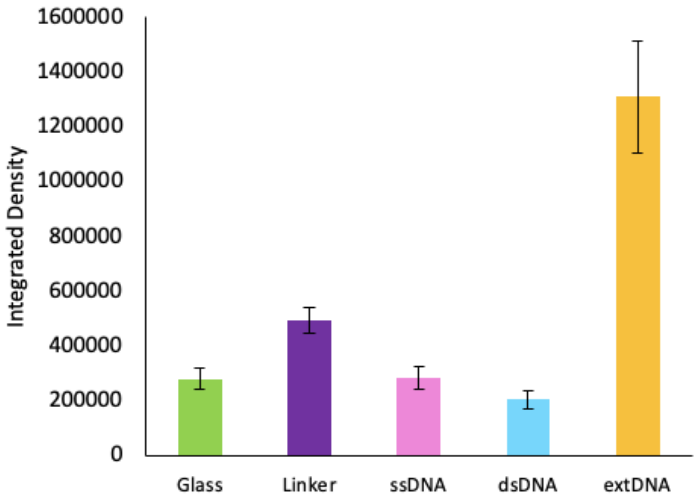


Figure 4.2. *Integrated density measured after each modification step of a $[GATC]_5/[GATC]_5$ surface.*

Little variation was observed in the measured PG fluorescence between the dsDNA and ssDNA surfaces since the dye is proven to interact with ssDNA.[16] Furthermore, it was difficult to determine an increase in fluorescence on intercalation to the dsDNA surface due to the short nature of the oligoseed, 20 bp. Therefore, the integrated density values were within error for these surfaces. It was clear that there was an 84 % increase in fluorescence on comparison of the short dsDNA surface to the extDNA brush surface. This large enhancement in fluorescence for the extDNA brush surface indicated that the Whitfield enzymatic extension was successful since the fluorescence increased due to the presence of more PG intercalation sites per molecule. The observed fluorescence density values for the glass and linker were due to interaction of the PG with the glass, however the values were low enough to be regarded as negligible since the increase in fluorescence was so large for the extDNA brush surface.

Further surface characterisation was performed by contact angle measurement to monitor the change in the surface after each modification step; clean glass, linker, ssDNA, dsDNA and ext DNA, **Figure 4.3**.

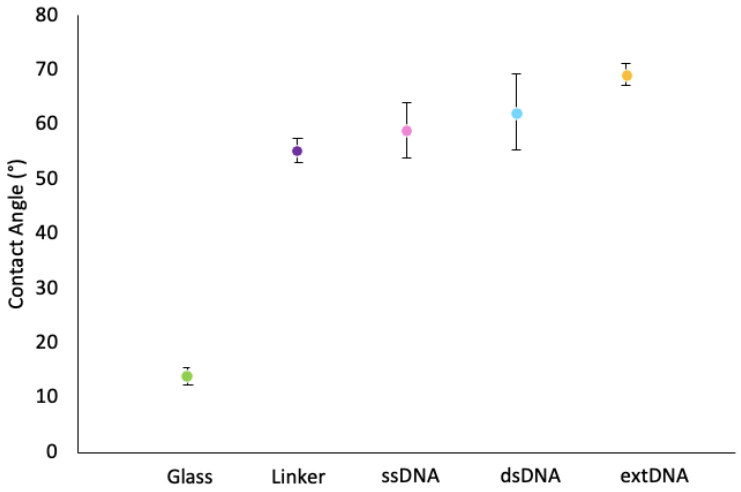


***Figure 4.3**. Contact angle of [GATC]$_5$/[GATC]$_5$ surface after each modification step.*

The contact angle of the clean glass surface was 14° which showed that the surface was hydrophilic (< 90°). After linker modification, the contact angle significantly increased to 55° and was more hydrophobic than the clean glass. It was also indicative of successful linker deposition as the surface morphology had changed. The contact angle only slightly increased after the immobilisation of ssDNA (59°) and hybridisation to generate dsDNA (62°). The immobilised oligoseed was short (20 bp) and did not induce much change on the surface. However, the error values of the ssDNA and dsDNA surfaces were much greater than the other samples which suggested the DNA attachment to the surface was not uniform. Once the brush surface was generated, the contact angle increased to 69° due to the hydrophobic nature of the DNA bases. The length of the extDNA was past that of the DNA persistence length (50 nm, 148 bp),[4] meaning the DNA was less rigid and more flexible than the dsDNA so there could be more bases in contact with the water. This would create a hydrophobic monolayer and therefore, cause an increase in contact angle.

Previous results confirmed that the surface modification and Whitfield enzymatic extension was successful,[3] however the number of heat-cool cycles was investigated to understand if it was possible to distinguish between the different DNA base pair lengths produced after 5, 10, 15 and 20 heat-cool cycles, **Figure 4.4**.

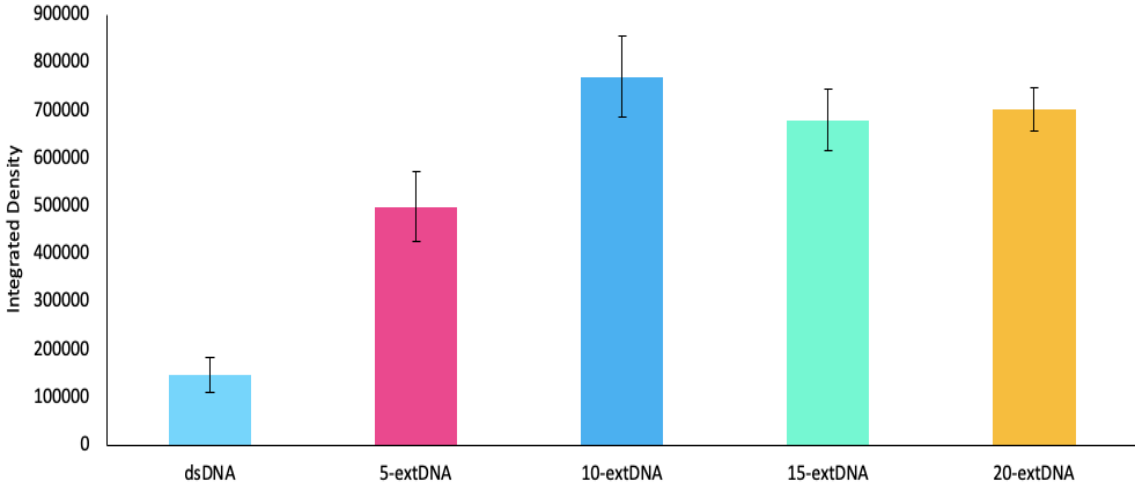


Figure 4.4. *Integrated Density of dsDNA surfaces and surfaces after 5, 10, 15 and 20 heat-cool cycles enzymatic extension.*

After the Whitfield enzymatic extension, it was clear that all heat-cool cycle lengths investigated produced brush surfaces with enhanced fluorescence density when compared with the dsDNA result. As shown in **Chapters 2 and 3**, the DNA length increased with respect to the number of heat-cool cycles and it was expected that the fluorescence density would increase in the same pattern. However, it was evident that the pattern was only upheld for 5 and 10 heat-cool cycles, as the fluorescence density decreased by 9-12 % for the 15 and 20 heat-cool cycles, although these values remained 80 % higher in comparison to the dsDNA. The variation between the samples may have been due to the measurements being recorded on several sample chips for each cycle set. Each of the chips will have modified to a different extent during linker modification and DNA attachment, so the surface was not uniform across all sample chips prepared. This will have affected the outcome of the Whitfield enzymatic extension as each chip will have extended differently.

Although it was clear from the $[GATC]_5/[GATC]_5$ Whitfield solution extension,[3] that 20 heat-cool cycles produced the longest extDNA length and were identified as the best conditions for surface extension, **Figure 4.4** showed that 5 heat-cool cycles distinguished the difference between dsDNA and extDNA well. Since, enhanced results were obtained 15 heat-cool cycles earlier than expected, the point at which no further surface modification occurs following the Whitfield extension was investigated. The contact angle was measured after 1, 2, 3, 4, 5, 10, 15 and 20 heat-cool cycles, **Figure 4.5**.

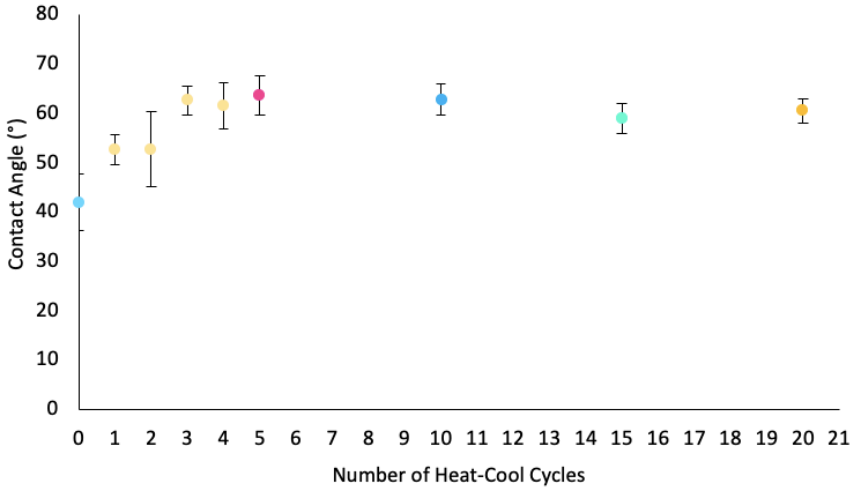


Figure 4.5. *Contact angle after increasing number of heat-cool cycles; 0, 1, 2, 3, 4, 5, 10, 15 and 20 cycles.*

The contact angle slowly rose as the number of heat-cool cycles increased from 0-4 cycles (yellow dots). After 5 heat-cool cycles, it was expected that the contact angle would continue to increase, however it was found that for 5-20 heat-cool cycles it remained similar. These heat-cool cycles were all within the error bar range of one another which was due to the extDNA being longer than the DNA persistence length. This result suggests that the DNA was very flexible and could collapse onto the surface due to instability of the DNA to remain upright, as described by Bracha *et al.*[4] Due to the DNA folding, the contact angle would not vary much after the persistence length had been reached since the surface would be of similar structure regardless of the DNA length.

Overall, the $[GATC]_5/[GATC]_5$ DNA oligoseed successfully proved that long DNA brush surfaces could be obtained by growing the DNA off the modified surface. This allowed for the distinction between short dsDNA and long extDNA surfaces and demonstrated that by increasing the number of base pairs the fluorescence could be enhanced which is a key requirement in the development of a DNA bio-sensor device.

4.2.2 Surface Extension – BAT25 Sequence

Microsatellite instability (MSI) is generated from the reduction in DNA mismatch repair (MMR) due to abnormal function. Cells which have irregular functioning MMR cannot correct errors and cause deletions and insertions which affect the allele length.[2,18–20] MSI markers are found in patients with tumours belonging to hereditary colorectal cancers.[21] The MSI structure involves non-coding nucleotide repeats which differ in allele length when comparing patient tumour cells to normal DNA.[22] Identifying the MSI present

indicates the tumour type and predicts the patient's chemotherapeutic response.[23] BAT25 is the most common mononucleotide repeat for MSI identification[24], and consists of a poly(T) repeat unit which involves significant base deletions in all tumours showing MSIs. Rapid, sensitive and reproducible MSI detection is essential for patient diagnosis and treatment.

The Bethesda panel[25] is the current MSI recognition method and consists of screening for five MSI markers on a specially designed panel. The five screened markers consist of mononucleotide repeat markers (BAT25 and BAT26) and dinucleotide repeat markers (D2S123, D5S346 and D17S250).[24] When the patient has 2-5 mutated markers, they are diagnosed with having a high level of MSIs. Unfortunately, this test is often deemed unreliable due to inconsistencies in the specificity and sensitivity of the panel.[26]

A DNA sensor which would increase the sensitivity and specificity of the test could be achieved by surface fabrication using the method described in **Scheme 4.6**. The BAT25 mononucleotide repeat was used previously to determine the identification of the MSI by base mismatches.[3] The research undertaken in this section aims to increase the reproducibility of the method through dehybridisation of the extended DNA surface and rehybridisation of different length target MSI DNA oligos.

The surface fabrication and extension were performed with the immobilisation of the probe oligo (5'-AmC_6-GAAAAAAAAAACGAAAAAAAAAAC-3') and hybridisation of the target oligo (5'-GTTTTTTTTTTCGTTTTTTTTTTC-3'). The dsDNA surface was extended for 20 heat-cool cycles using the Whitfield enzymatic extension and initially characterised by contact angle measurements, **Figure 4.6**.

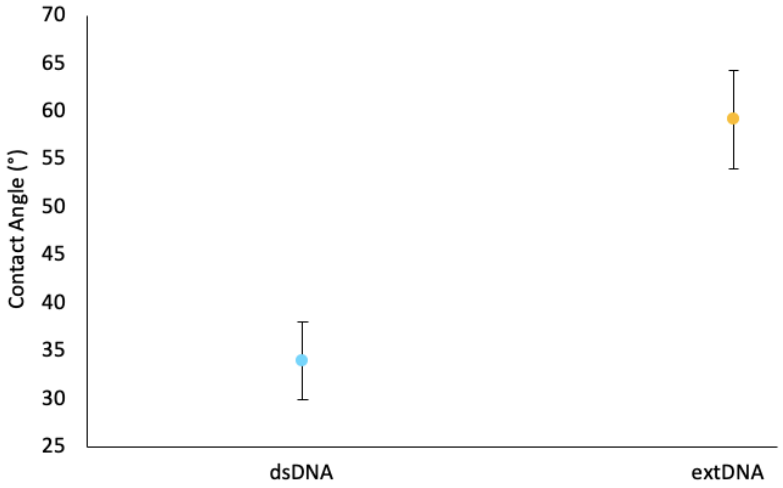


Figure 4.6. *Contact angle after dsDNA modification and 20 heat-cool cycle Whitfield extension of the surface.*

The contact angle increased from 34° on the dsDNA surface to 59° on the extDNA surface. Enzymatic extension of the surface was successful due to the increase in hydrophobicity from the greater number of bases present, indicating how the DNA length affected the structure of the surface, as further supported by the corresponding fluorescence measurements, **Figure 4.7**.

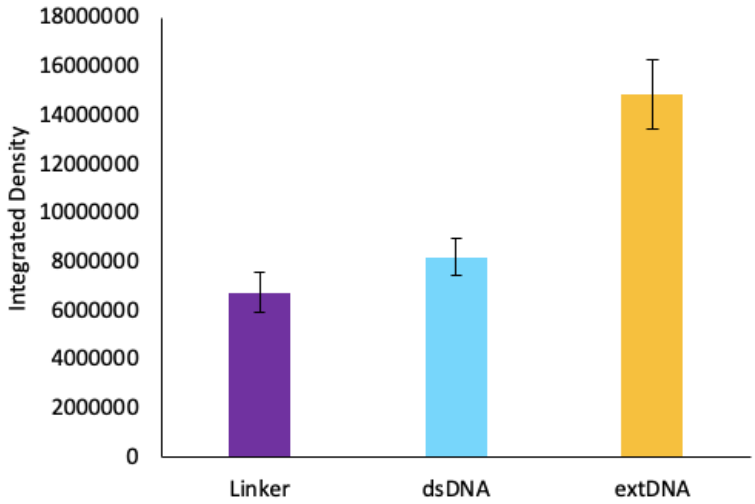


Figure 4.7. *Integrated density of linker modification, dsDNA surface immobilisation and 20 heat-cool cycle Whitfield extension.*

Enhancement of the fluorescent signal supported the contact angle measurements, as the dsDNA integrated density was almost 50 % lower than that of the extDNA. The fluorescence signal only slightly increased from the linker modification to dsDNA modification due to the short nature of the DNA and the low output signal it produced from the limited number of intercalation sites. The surfaces had extended in DNA length

well and good fluorescent enhancement of the surface was achieved. Therefore, the BAT25 surfaces were good candidates for the dehybridisation/rehybridisation reproducibility studies.

As mentioned previously, this section is focussed on the reproducibility of MSI detection by base mismatch rehybridisation. For rehybridisation to occur, the extDNA surfaces were dehybridised at 95 °C for 10 minutes to remove the long ssDNA. The ss-extDNA surface was re-heated in a solution of complement target DNA at 75 °C for 10 minutes and allowed to cool slowly to room temperature to enable hybridisation to the DNA on the surface. The fluorescence was measured after the dehybridisation and rehybridisation steps, **Figure 4.8**.

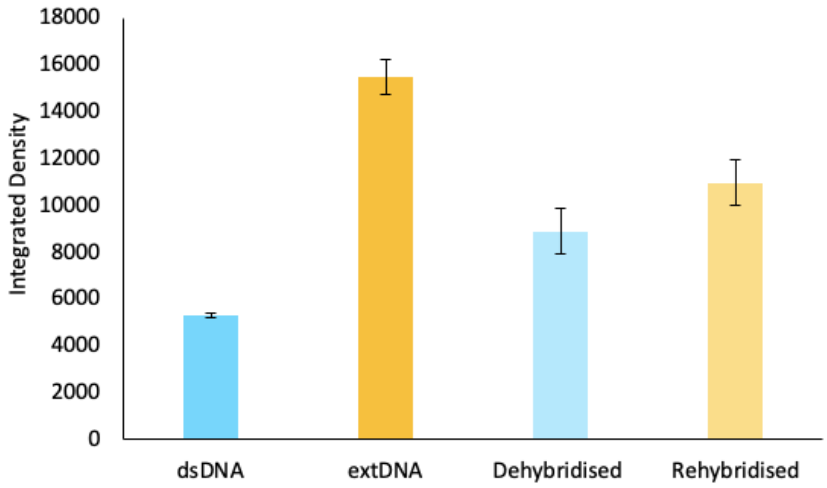


Figure 4.8. *Dehybridisation of extDNA surface to a ss-extDNA surface and rehybridisation of complement DNA to reform the hybridised structure.*

Successful surface extension of the BAT25 oligoseed was noted due to the significant increase (~ 190 %) between the dsDNA and the extDNA. After dehybridisation, the fluorescent signal decreased by 40 %, which indicated partial removal of the ssDNA from the surface. Although, some of the dehybridised fluorescence may be attributed to the signal from the weak binding affinity PG has to ssDNA.[16] The rehybridised surface only increased by ~ 20 % from the dehybridised surface. This suggested that the annealing conditions were not optimised as it was thought the signal should enhance similar to that of the extDNA on complementary target rehybridisation.

Furthermore, it was expected that the dehybridised surfaces would have decreased in integrated density to values below that of the dsDNA since the surface now consisted of ssDNA. The high fluorescence value obtained for the dehybridised surface may be due to

the self-annealing of A-A base mismatches on the ss-extDNA surface. The increased flexibility of the surface on dehybridisation of the extDNA to ssDNA, would have allowed for the DNA to fold and self-anneal on cooling. It was not anticipated that these base mismatches would enhance the signal significantly, so it is more likely a co-factor to the surface not fully dehybridising in the time given. In order to obtain better dehybridisation, optimisation of the heating time at 95 °C was considered, **Figure 4.9**.

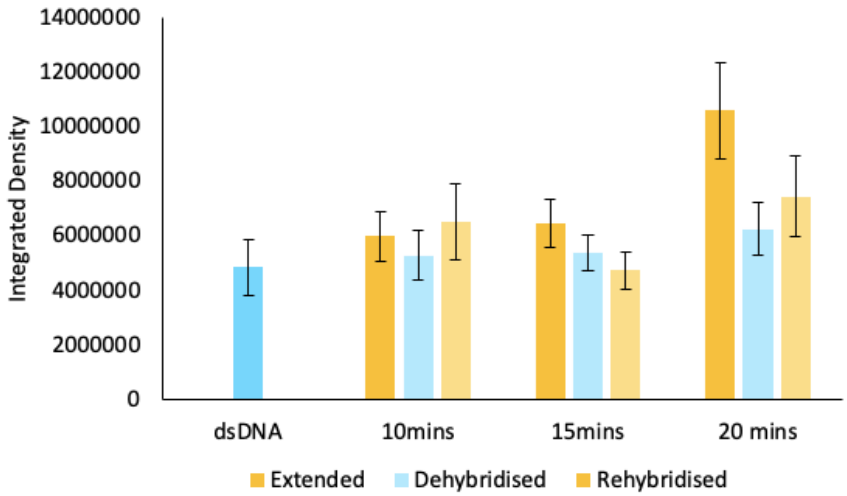


Figure 4.9. *Optimisation of heating time for dehybridisation at 95 °C. Fluorescence measured and integrated density calculated after 20, 30 and 40 minutes.*

Firstly, it was clear that the extDNA surfaces used for 20 and 30 minutes dehybridisation did not extend well from the dsDNA surface, with fluorescence enhancement increases of just 23 and 33 % observed, respectively. However, there was an 118 % enhancement in fluorescence for 40 minutes. Unless the surface showed no fluorescent enhancement, *i.e.* no DNA extension occurred, the degree of extension should not have affected the dehybridisation. The highest quantity of dehybridisation was observed for the 40 minutes heating time, with a 40 % reduction in integrated density recorded. This was potentially due to the greater surface extension, although it was more likely due to the increased time given for the long DNA strands to unwind from the duplex structure. Dehybridisation for 20 and 30 minutes afforded reductions of 12 and 16 %, respectively. Hence, dehybridisation was undertaken for 40 minutes at 95 °C in future experiments.

The rehybridisation of the surfaces in **Figure 4.9** could not provide definitive detection of the binding of the 26 bp complement target oligo. This was due to the increase in integrated density only being slightly higher, on most occasions, than the dehybridised surface, ~ 15 %. To improve the probability of detecting the target DNA rehybridisation,

a study focussed around the Tm of the target oligo was undertaken. It was noticed that the annealing temperature of 75 °C used for rehybridisation was significantly higher than that of the calculated oligo Tm by OligoCalc, 45 °C. Therefore, the annealing temperature was reduced from 75 °C to 45 °C and the fluorescence on rehybridisation was measured, **Figure 4.10**.

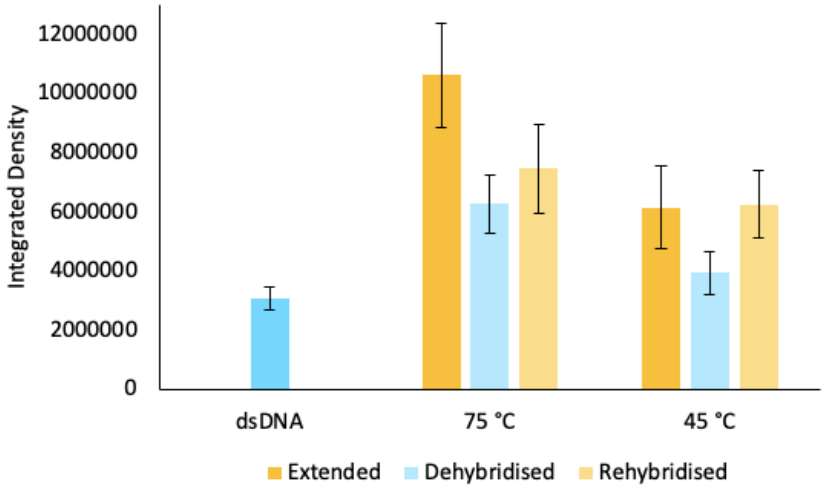


Figure 4.10. *Optimisation of rehybridisation temperature. Fluorescence measured, and integrated density calculated after surfaces were slowly cooled to room temperature from 10 minutes at 75 °C or 45 °C.*

Surface extension of the dsDNA surface was deemed successful, as an increased fluorescent signal was observed for both temperatures, up to 118 % above that of the dsDNA signal. The dehybridisation of the surfaces was still not measured below that of the initial dsDNA sample. However, the rehybridisation of the sample annealed at 45 °C exhibited re-enhancement of the surface by the complement target DNA within ~ 2 % of the initial extDNA surface. Hence, an annealing temperature of 45 °C was used in future target DNA annealing studies.

The optimised conditions were subsequently used to dehybridise and rehybridise the extDNA surfaces multiple times, with the resulting fluorescence measured after each step to monitor the reproducibility of the newly developed methods, **Figure 4.11**.

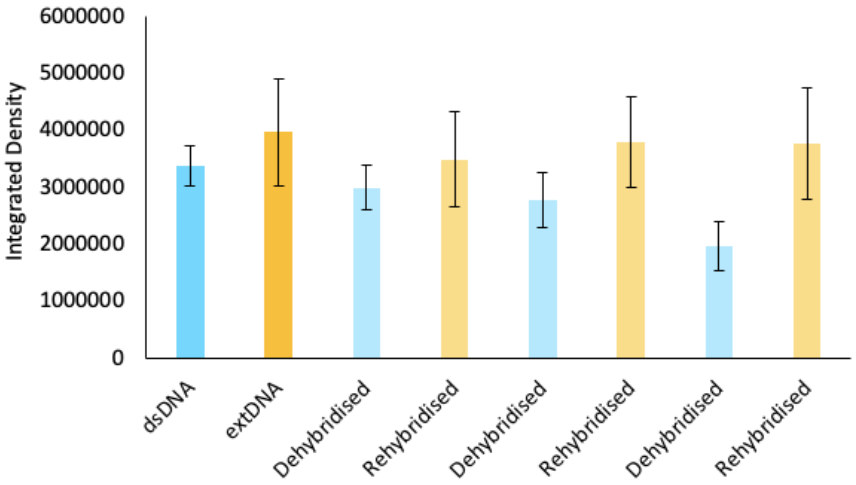


Figure 4.11. *Fabricated extDNA surfaces dehybridised at 95 °C for 40 minutes and rehybridised with complement target DNA at 45 °C for 10 minutes before cooled to room temperature. The fluorescence was measured after each dehybridisation and rehybridisation step.*

Unlike previous attempts, all of the dehybridisation steps afforded fluorescence results lower than the dsDNA, with the final dehybridisation integrated density reduced by 42 %. The multiple rehybridisation steps also proved successful, as all fluorescence values for the complement DNA oligo were within error of the enhanced extDNA measurements. The final rehybridisation was within 5 % of the initial fluorescence measurement of the extDNA, hence proving the reproducibility of the surface.

After proving the reproducibility of the sample by multiple dehybridisations and rehybridisations, the effectiveness of the fabricated DNA sensor was analysed via the identification of several target sequences. Before applying the mutated sequences for MSI detection, a random 20 bp non-complementary target sequence was used (5'-CAGTCGTTCCCAGTCGTTCC-3') for rehybridisation, **Figure 4.12**.

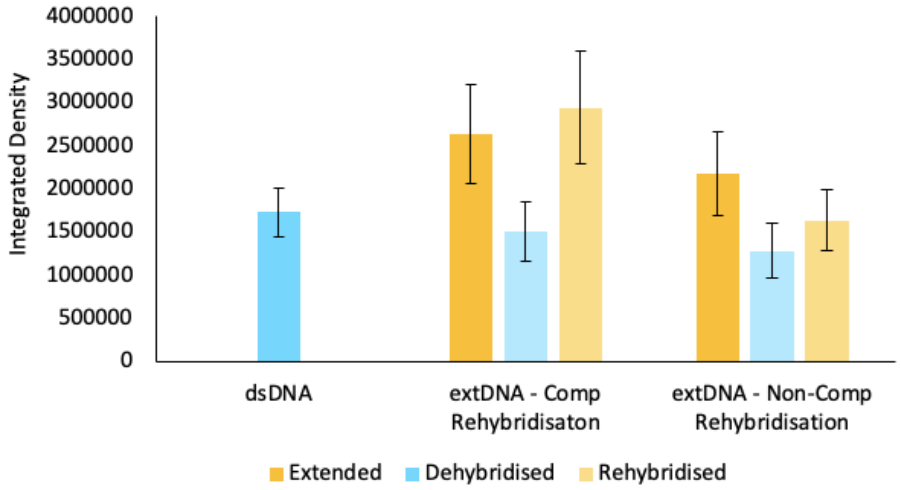


Figure 4.12. *Modified BAT25 surfaces rehybridised with target DNA; BAT-25 complement and a non-complementary sequence.*

The complementary target DNA strand rehybridised to a sufficient standard, with the fluorescence returning to within error of the initial measured level. The 10 % increase in fluorescence between the extDNA and the rehybridised surface may be due to the reorganisation of the surface during heating, which allowed for more probe strands to become available for binding.[27] The non-complementary sequence was 75 % rehybridised in comparison with the initial extDNA integrated density. This was unexpectedly high as it was assumed that the result would be closer to that of the dehybridised surface. However, it should be noted that the error bars overlap so a clear indication of non-specific target DNA binding could not be determined.

For colorectal cancer MSI detection, the differentiation between single, double and triple base mismatches on target strand hybridisation was required to be identified by the fabricated DNA sensing platform. The dehybridised extDNA surfaces were rehybridised with mutated target DNA oligos which contained one (5-'GTTTTTTTTTTC-3'), two (5'-GTTTTTTTTTC-3') and three (5'-GTTTTTTTTC-3') T base deletions, **Figure 4.13**.

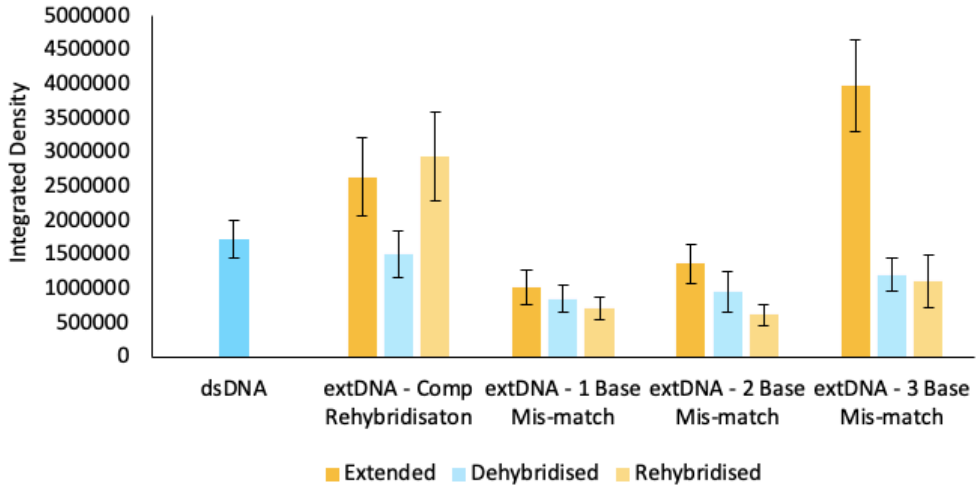


Figure 4.13. *Surfaces rehybridised in DNA sequences with one, two and three base mismatches.*

All three mutated target DNA sequences did not fully rehybridise with the immobilised probe strand to the extent of the complementary DNA, recorded at 125 % relative to the initial extDNA surface. The single base mismatch rehybridised to 70 % of the extDNA surface, which was expected due to the similarity of the complementary and the one base mutated sequence. The double base mismatch rehybridisation was 45 % of the extDNA surface. This value was lower than that achieved by the single base mismatch, however it rehybridised more than the triple base mismatch target, 30%. It was evident that the sensor could distinguish between the different mutated DNA target sequences and aid in the identification of the MSI present. This was despite the varying degrees of surface extension evident in the extDNA surfaces due to non-uniform linker deposition. Although further reproducibility testing was required to obtain certainty in the detection method of base mutations, the results were positive in the journey to develop a DNA sensor for small genetic mutations such as MSIs.

4.2.3 Surface Extension – CFTR Sequence

The cystic fibrosis transmembrane conductance regulator (CFTR gene) is another disease where biosensing by specific sequence differentiation would be beneficial by offering early disease detection. The ABC transporter class ion channel which the CFTR gene codes for, allows for chloride and thiocyanate ions to migrate through the epithelial cell membranes in the lungs, pancreas and other organs.[28–31] CF is most commonly the result of a three base gene mutation, where CTT bases are deleted in the CFTR gene. This induces the loss of the phenylalanine amino acid and encourages protein mis-folding which distorts the shape causing the loss of ion-channel function.[1,29,32,33]

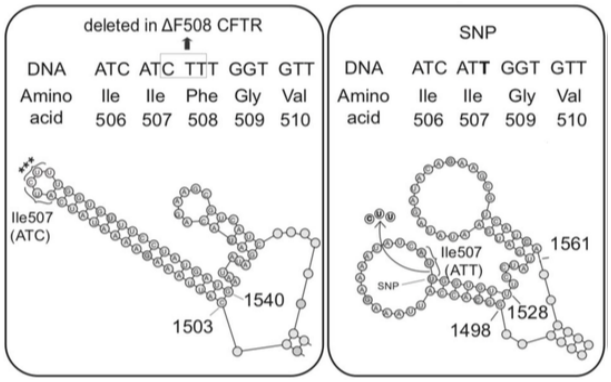


Figure 4.14. *CFTR mutation by the deletion of phenylalanine amino acid (CTT DNA bases) causing gene shape distortion.*[33]

The consequences to the patient of ion-channel malfunction are increased respiratory infections due to mucus thickening in the lungs and pancreatic blockages leading to malnutrition and diabetes.[31] As these complications give rise to chronic disability and lowered life expectancy, the ability to detect the gene deletion early would enable immediate patient treatment., decreasing the effects of the disease and increasing the patient's quality of life.

For the generation of a fluorescence enhanced DNA sensor device to detect the CFTR gene mutation, the probe sequence (5'-AmC_6-GCATCTTTCGGCATCTTTCG-3') and target sequence (5'-CGAAAGATGCCGAAAGATGC-3') were used to fabricate the surface as per **Scheme 4.6**. Since it was already proven that the sequence extended off of the surface to produce an enhanced fluorescence signal,[3] the Whitfield enzymatic extension was used to monitor the fluorescence enhancement with increasing cycle number, **Figure 4.15**.

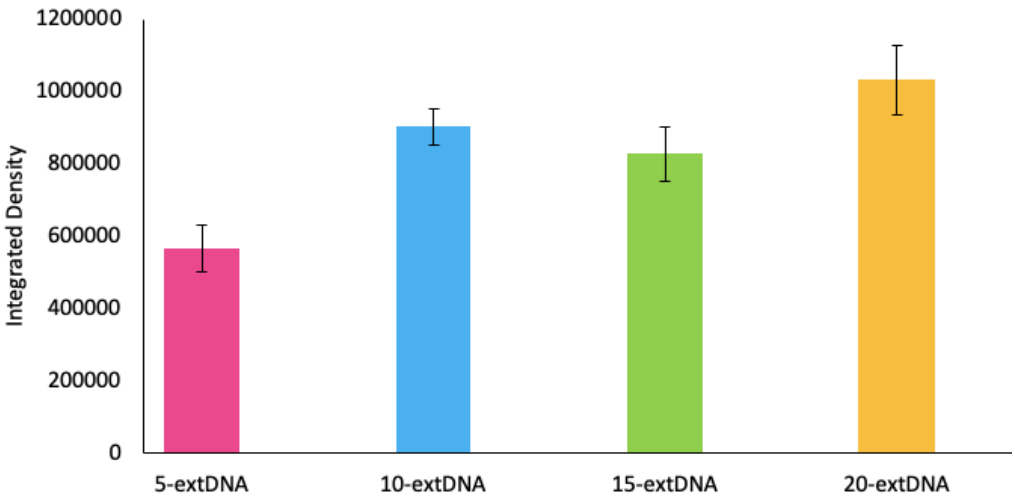
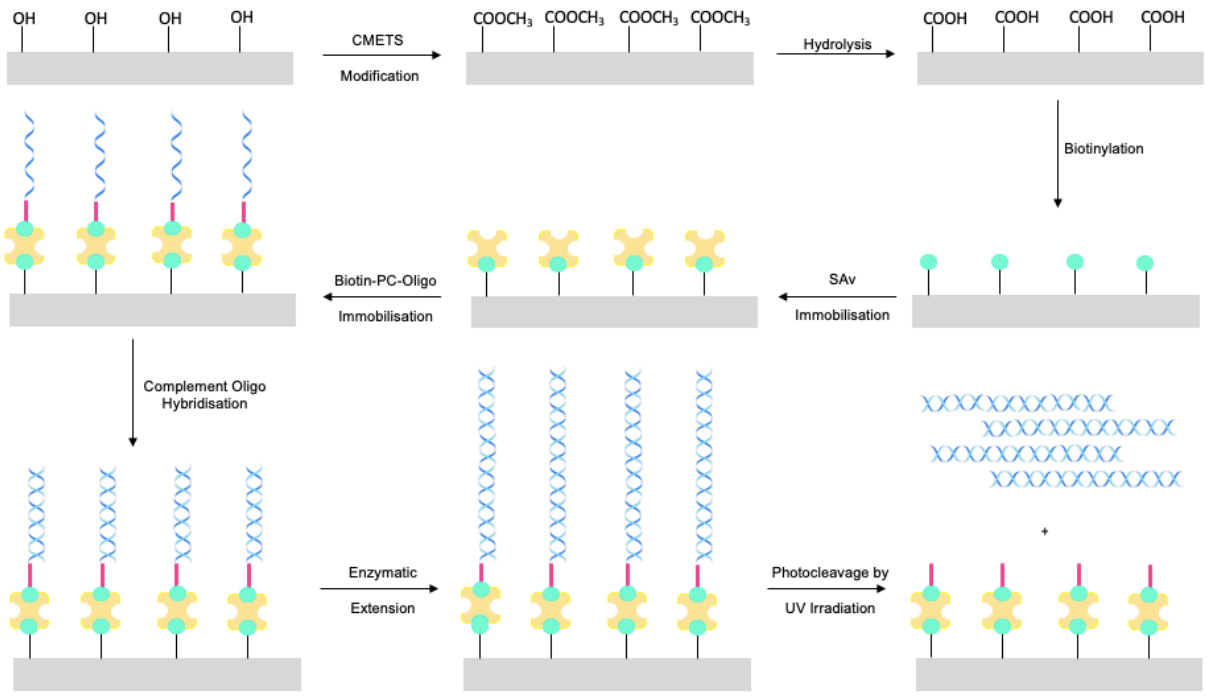


Figure 4.15. *Integrated density of CFTR surface after 5, 10, 15 and 20 heat-cool cycles.*

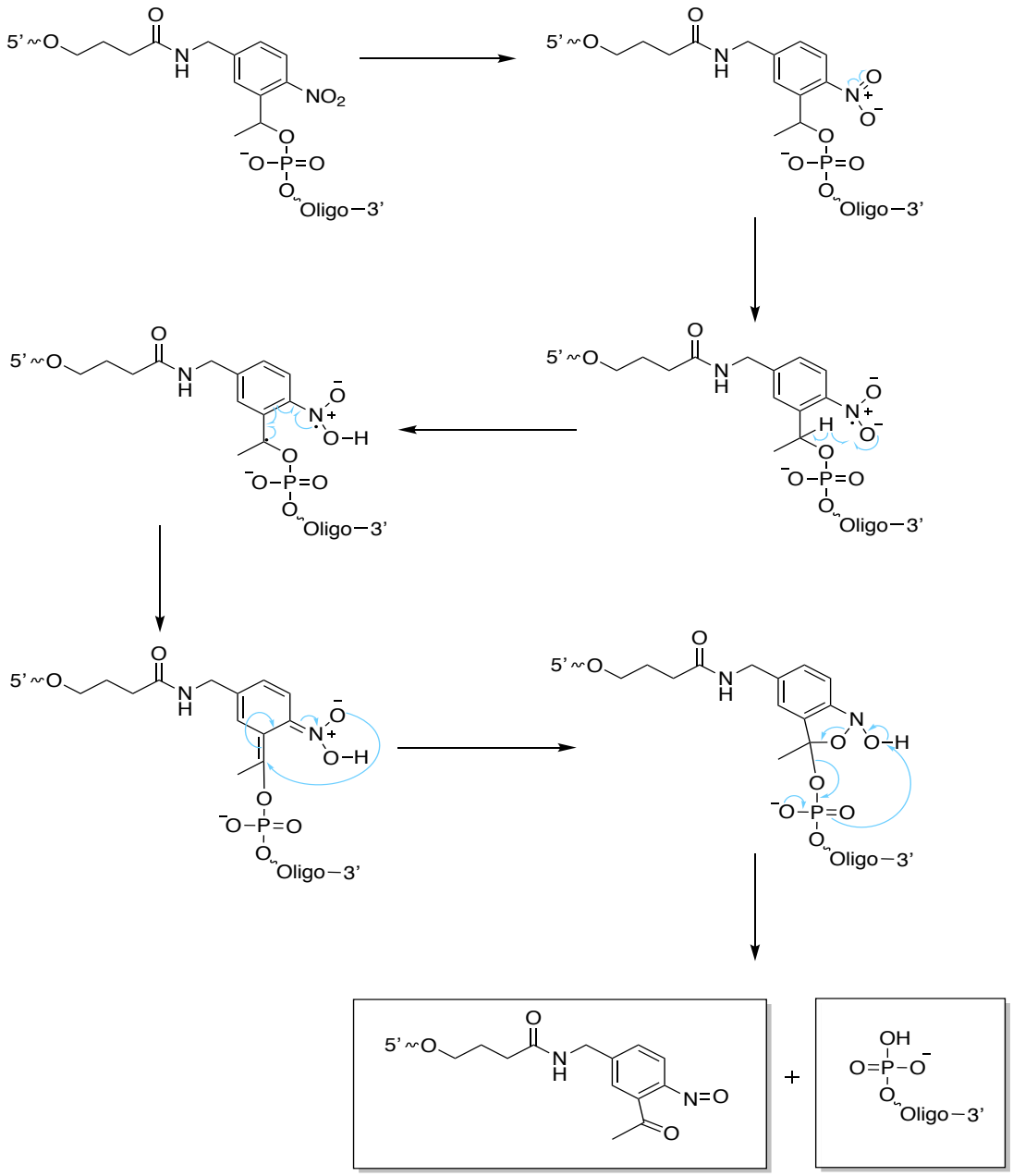
The fluorescence signal increased relative to the number of heat-cool cycles and confirmed the increase in intercalation sites due to continued DNA extension. However, it was clear that the error bars for 10, 15 and 20 heat-cool cycles overlapped. This was due to the unavailability of a significant number of binding sites after 5 heat-cool cycles, as the DNA had grown beyond that of the persistence length. The structure of the extDNA on the surface would collapse to the Bracha model,[4] as discussed in the introduction to this chapter, where the DNA was folded over on the surface, rather than standing upright and therefore, steric hindrance of the flat DNA on top of one another would obstruct intercalation of the PG. Although this result showed that the number of heat-cool cycles does affect the fluorescent signal enhancement, further research could be undertaken to further increase the signal enhancement by prohibiting the DNA from collapsing onto the surface.

Previous research by Lunn[3] focussed on the determination of the base mutations and the possible use of the device for CFTR disease detection. However, the understanding of the extDNA length on the surface was not investigated. The length of extDNA on the surface is crucial to the enhancement of the fluorescence signal and so this section focusses on determining the extDNA length after surface extension. This research was performed in collaboration with Prof Ijiro (Hokkaido University, Sapporo, Japan) to utilise surface fabrication techniques, in particular via photo-irradiation to create a DNA pattern by the deactivation of a SAv surface.[11,34] It was intended that the UV treatment of the surface was combined with research reported by Agasti *et al.*[35] which demonstrated the release of DNA oligos by photocleavable linkage, **Scheme 4.7.** This would enable the extDNA length to be characterised by agarose gel, UV-Vis spectroscopy and AFM.



Scheme 4.7. *Surface fabricated by streptavidin-biotin modification and biotin-photocleavable DNA hybridisation. Photocleavable linker indicated by pink line. Surface irradiated with UV to cleave the linker and release the DNA from the surface.*

The glass surface was modified with 2-(carbomethoxy)ethyltrichlorosilane (CMETS) and hydrolysed with hydrochloric acid. Biotinylation of the surface was achieved via amine coupling and SAv immobilisation. The DNA probe strand contained a photocleavable linker, which was inserted after the 5' amine modification for DNA oligo attachment to the surface and before the start of the DNA sequence. The ssDNA target probe was hybridised to generate the dsDNA oligoseed. The surface was irradiated with 600 mW/cm^2 UV light at 365 nm for between 0 and 20 minutes, with the resulting fluorescence measured. The mechanism of photolysis is outlined in **Scheme 4.8**.



Scheme 4.8. *Photolysis mechanism of the photocleavable linker contained within the DNA oligo.*[36]

The SAv-biotin approach used by the Ijiro group was never enzymatically reacted above 37 °C and consequently, the stability of the SAv-biotin interaction at 95 °C was untested. Therefore, an initial experiment to check the stability of the SAv-biotin DNA modification was performed. A DNA surface was prepared, as per **Scheme 4.7**, with a biotin-photocleavable DNA probe oligo which contained an Alexa488 fluorophore on the 3' end and was hybridised with the complement target oligo. The fluorescence of the freshly modified dsDNA surface was recorded. The sample was placed in an Eppendorf which contained 200 µL of 10 mM Tris buffer and was inserted into a hot-block pre-heated to 95 °C and left for 10 minutes. These conditions were the total amount of time the sample would see 95 °C during 20 heat-cool cycles of the Whitfield enzymatic extension.

The fluorescence of the surface was re-measured on removal from the hot-block, **Figure 4.16.**

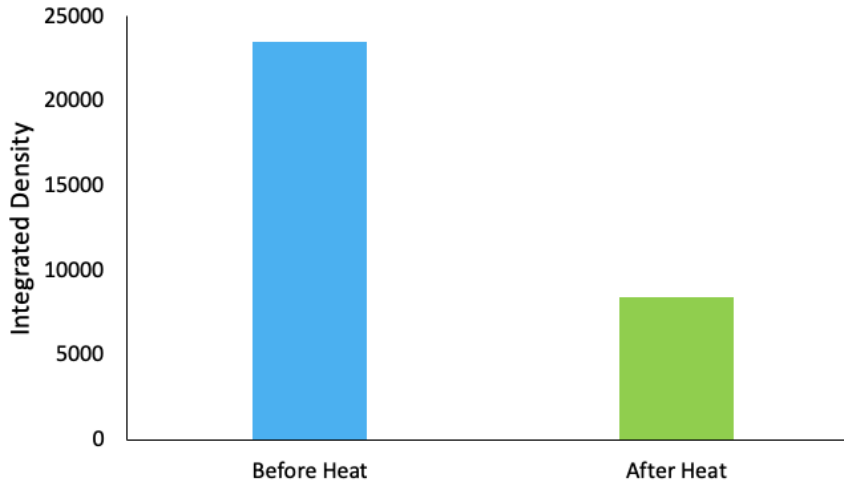


Figure 4.16. *Integrated density of SAv-biotin dsDNA + Alexa647 modified glass surface before and after heating to 95 °C for 10 minutes to check biotin stability.*

It was found that the fluorescence of the surface decreased by 70 % after the 10 minutes of heating, so the biotin linkage was not stable at high temperatures and could not be used to grow the DNA by the Whitfield enzymatic extension.

Before photocleavable CFTR DNA immobilisation on the APEGDMES modified surface, the DNA immobilisation and hybridisation were performed with the standard CFTR probe and target oligo strands and the dsDNA surface was extended for 20 heat-cool cycles. The surface was irradiated for various times with UV at 15 mW power and washed for 15 minutes in a jitterbug water wash. This experiment would determine the effect of UV irradiation on the intercalated PG dye on the surface, **Figure 4.17**.

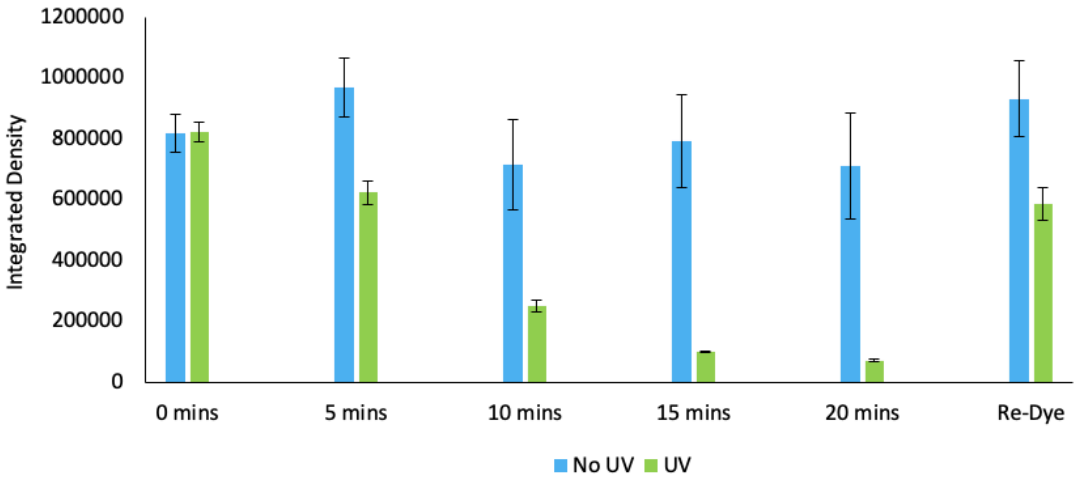


Figure 4.17. *Integrated density of non-photocleavable CFTR DNA after UV irradiation for 5, 10, 15 and 20 minutes, compared to surfaces which were not UV treated. Fluorescence measured after 15 minutes jitterbug water wash.*

The UV-irradiated non-photocleavable surface showed a loss of fluorescence as the irradiation time increased. Since the DNA on the surface did not include a photocleavable linker, it was not expected that the fluorescence would decrease by 65 % after 20 minutes of UV irradiation. Furthermore, the surfaces which were not treated with UV irradiation produced fluorescence results which were within error as expected due to no induced change on the surface. The re-dying of the surface enhanced the fluorescence back to its original level which confirmed that little to no DNA was removed during this process. Therefore, it was concluded that the loss of fluorescence was due to the high-powered UV treatment of the surface which caused degradation of the PG dye.

Fresh surfaces were modified with the APEGDMES linker and a photocleavable linker was inserted between the 5' amine and the DNA sequence on the probe oligo (5'-AmC_6-iSpPC-GCATCTTTCGGCATCTTTCG-3', IDT, Singapore). The photocleavable surface was extended for 20 heat-cool cycles and irradiated for various time periods with UV at 15 mW power. The irradiated surfaces were washed for 15 minutes in a jitterbug water wash and the fluorescence measured on a BioRad molecular imager, **Figure 4.18**.

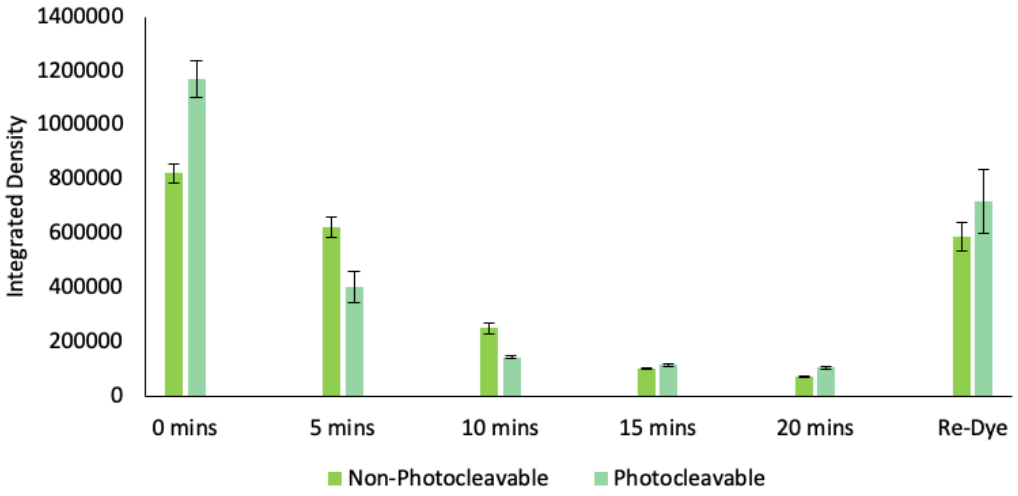


Figure 4.18. *Fluorescence density of photocleavable CFTR DNA compared to non-photocleavable CFTR DNA after UV irradiation for 5,10, 15 and 20 minutes. Fluorescence measured after 15 minutes jitterbug water wash.*

The set of photocleavable DNA surfaces which underwent UV irradiation for 5, 10, 15 and 20 minutes showed that the fluorescence signal decreased by 65 % after 5 minutes of exposure, however the non-photocleavable DNA surfaces had only decreased by 41 % in the same time frame. After 20 minutes of UV irradiation, the photocleavable surface had reduced by 91 % compared to the 0 minutes sample. In comparison, the non-photocleavable DNA had only decreased by 65 %. Furthermore, it was evident that DNA still remained on the surface as after 20 minutes UV irradiation and re-dying the surface, the photocleavable DNA fluorescent signal increased by 90 % from the 20 minutes sample and illustrated that DNA intercalation was still possible. Some DNA had been removed since the intensity was less than the initial measurement and suggested that there were two effects occurring on the photocleavable DNA surface. Firstly, as identified from **Figure 4.17**, the PG degradation due to the high-power UV treatment and secondly, that some of the DNA was cleaved from the surface. This explained the sharper decrease in fluorescence observed in the photocleavable DNA surfaces compared to those which were fabricated with the non-photocleavable DNA.

Aside from fluorescence measurements, the washes from the UV-irradiated photocleavable DNA surfaces were concentrated and characterised by UV-Vis spectroscopy and agarose gel electrophoresis, **Figure 4.19**.

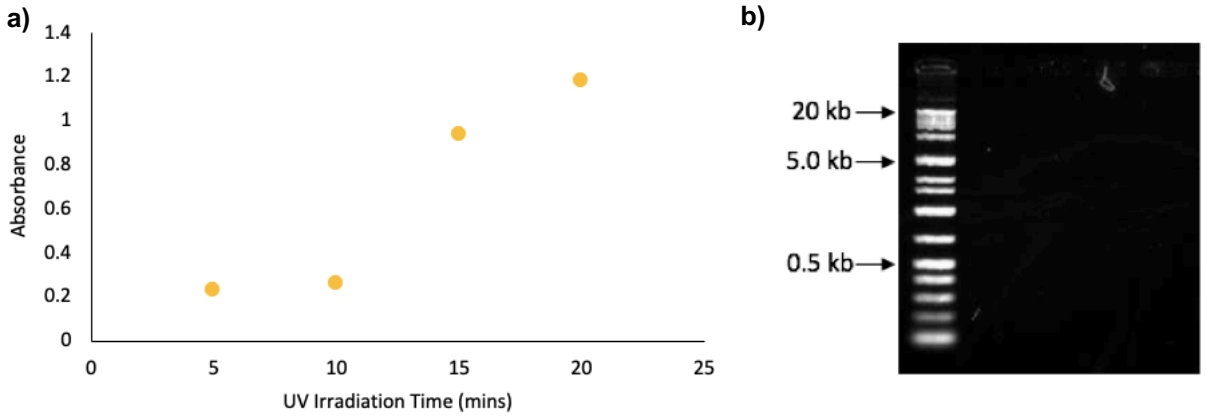


Figure 4.19. *Characterisation of photocleavable CFTR DNA removed from the surface after 5, 10, 15 and 20 minutes UV irradiation time.* ***a)*** *Absorbance values recorded at 260 nm* ***b)*** *1 % agarose gel electrophoresis run at 100 V.*

The 260 nm absorbance values indicated that DNA was recovered from the surface after photocleavage of the linker by UV irradiation. It was clear that the longer the surface was irradiated for, the higher the concentration of DNA recovered. After 5 minutes the DNA concentration was 11.6 ng/µL and after 20 minutes the concentration rose to 58.9 ng/µL. This contradicted the findings from the fluorescence study which concluded that the DNA had not cleaved well from the surface. However, the 1 % agarose gel displayed no evidence of the presence of DNA after 5, 10, 15 or 20 minutes. This suggested that the DNA removed from the surface may have been in single base form rather than single stranded or duplex form due to sequence degradation from the intense UV light irradiation.

Although the results discussed in this section did not yield a DNA length, the use of a photocleavable linker within the probe DNA oligo is still a possible route for length determination after surface extension. However, more work is required on the optimisation of the removal of the DNA by UV methods.

4.3 Conclusions

The Whitfield enzymatic extension method was further explored with regards to surface immobilised DNA for sensing applications. The main aims of this chapter were to address the issues associated with reproducibility of the target DNA detection and to characterise DNA lengths post surface extension. This was achieved by using three different sequences; one theoretical sequence, $[GATC]_5/[GATC]_5$, and two real-life mutated sequences, BAT25 and CFTR.

$[GATC]_5/[GATC]_5$ probe and target sequences were used as a proof of concept oligoseed for the production of long DNA brush surfaces by enzymatic extension of the immobilised oligoseed off the surface. Fluorescence of the intercalating dye, PG, was measured and the 84 % enhancement in fluorescence allowed for distinction between dsDNA and extDNA surfaces. This enhancement was due to the increase in the number of binding sites after Whitfield enzymatic extension which enabled more PG dye to intercalate and increase the fluorescence output signal.

The determination of the use of the fabricated DNA sensor to identify small MSIs in colorectal cancer diagnosis was explored by immobilising and extending the surface with the most common MSI in colorectal cancer, BAT25. The BAT25 extDNA surfaces were dehybridised to enable the rehybridisation of the target DNA sequences, however issues were identified with this process. Optimisation of the conditions concluded that dehybridisation for 40 minutes at 95 °C and rehybridisation at 45 °C, provided the best results. Real-life DNA mutations were rehybridised to the ss-ext-DNA surface to ensure the DNA sensor was fit to recognise small base deletions often associated with gene mutation and the production of MSIs. The results of these tests showed that a single base mutation rehybridised by 70 %, a double base mutation by 45 % and a triple base mutation by 30 % which indicated the ability of the surface to determine sequence differences and identify the MSI present.

Finally, the last part of this chapter considered obtaining a DNA length by photocleavage of the extDNA from the surface. Initially, the SAv-biotin surface modification was investigated, however, it was established that the surface could not withstand the high temperatures required during the heat-cool cycles in the Whitfield enzymatic extension. Therefore, the APEGDMES linker was used alongside a probe DNA strand which included a photocleavable linker between the 5’ amine and the DNA sequence. Photocleavage of

the surface was performed by UV irradiation and it was evident that there was degradation of the intercalated PG on comparison of non-photocleavable DNA and photocleavable DNA since in both cases there was a reduction in fluorescence. However, it was thought that some DNA was cleaved from the surface due to presence of a 260 nm peak in the UV-Vis trace from the collected surface washes after UV treatment.

Overall, the research presented in this chapter afforded positive results in the fabrication of a DNA sensor with the capability to determine gene mutations by enhanced fluorescence detection, although further work to fully optimise the sensor is required.

4.4 Experimental

4.4.1 Materials

Deoxynucleotide Preparation

Deoxynucleotides (**Table 4.1**) were purchased from Eurofins (Ebersberg, Germany).

Table 4.1. Purchased deoxynucleotides

Name	Origin	Sequence
Amine-GATC	Eurofins	5'-AmC_6-GATCGATCGATCGATCGATC-3'
GATC Comp	Eurofins	5'-GATCGATCGATCGATCGATC-3'
Amine-BAT25	Eurofins	5'-AmC_6-GAAAAAAAAAACGAAAAAAAAAAC-3'
BAT25 Comp	Eurofins	5' - GTTTTTTTTTTCGTTTTTTTTTTC-3'
BAT25 Non-Comp	Eurofins	5'-CAGTCGTTCCCAGTCGTTCC-3'
1 Base Mutation	Eurofins	5-'GTTTTTTTTTTC-3'
2 Base Mutation	Eurofins	5'-GTTTTTTTTTC-3'
3 Base Mutation	Eurofins	5'-GTTTTTTTTC-3'
Amine-CFTR	Eurofins	5'-AmC_6-GCATCTTTCGGCATCTTTCG-3'
CFTR Comp	Eurofins	5'-CGAAAGATGCCGAAAGATGC -3'
Amine-PC-CFTR	IDT DNA	5'-AmC_6-iSpPC-GCATCTTTCGGCATCTTTCG-3'

Expression and Purification of DNA Polymerases

Table 4.2. Description of DNA polymerases for enzymatic extension

DNA Polymerase	Properties	Source
Thermococococcus gorgonarius polymerase B (Pfu-Pol) Z3 exo-	Archael family B Polymerase Low fidelity variant with the 5'→3' exonuclease activity removed and alterations to the fingers domain	Gifted from Dr. Colette Whitfield (PDR)
Pyrococcus polymerase B Deep Vent exo-	Family B polymerase with 5'→3' exonuclease activity removed	New England BioLabs Inc.

4.4.2 Surface Preparation

Hokkaido University

Glass microscope slides were cut into chips of size 0.5 cm x 0.5 cm. The surfaces were cleaned in piranha solution for 1 hour and washed with nanopure water before the chips were dried with N_2. In a clean glass petri dish, the chips were placed under ultra-violet (UV) light for 1 hour.

Newcastle University

Glass microscope slides were cut into chips of size 0.5 cm x 0.5 cm. The surfaces were wiped with acetone, isopropanol (IPA) and nanopure water (H_2O). The chips were then placed in the sonicator and washed with acetone, IPA and nanopure water for a period of 15 minutes per solvent. In a glass petri, the chips were treated with O_2 plasma for 10 minutes.

4.4.3 APEGDMES Surface Modification

The pre-cleaned glass surfaces were immersed in APEGDMES/toluene solution (233 μM, 3 mL) preheated to 65 °C for 16 hours. The chips were washed with toluene, ethanol and nanopure H_2O three times sequentially and dried with nitrogen. Surfaces were placed under vacuum in a vacuum oven at 120 °C for 40 minutes.

4.4.4 CMETS Surface Modification

2 μL 2-(carbomethoxy)ethyltrichlorosilane (CMETS) was added to 20 mL toluene in a nitrogen environment and left for 1 hour. Surfaces were washed with acetone, methanol and nanopure H_2O, three times sequentially and dried with nitrogen. 6 M hydrochloric acid was poured on top of the surfaces in a glass petri dish and left overnight. The surfaces were washed three times with nanopure H_2O.

4.4.5 Streptavidin-biotin Modification

Biotin modification solution was made with 50 mM *N*-(3-dimethylaminopropyl)-*N'*-ethylcarbodiimide hydrochloride (EDC), 10 mM HEPES buffer and 2 mM EZ-link-amine-PEG_2-biotin (ThermoFisher, UK). 10 μL of the biotin solution was drop cast on to the surface and left for 1 hour. Surfaces were washed three times with 10 mM Tris-HCl (pH 7.6).

Streptavidin immobilisation solution was prepared with 10 mM Tris -HCl (pH 7.6) and 0.2 mg/mL streptavidin (Sigma Aldrich, UK). 10 μL of the streptavidin solution was drop cast onto the surface and left for 1 hour. The surfaces were washed three times with 10 mM Tris-HCl (pH 7.6).

4.4.6 DNA Immobilisation

Amine terminated DNA

Amino-tagged DNA probe solution (8 μL, 100 μM) in 10% acetic acid solution was drop cast onto the glass surfaces for 1 hour in a humid environment. $NaCNBH_3$ (8 μL, 16 μM) in 50 % methanol solution was deposited on top of the probe solution for a further 2 hours in a humid environment. The chips were washed with 0.5x phosphate buffered saline (PBS) and 3 mL nanopure H_2O to remove any excess DNA molecules. Surfaces were dried with nitrogen.

Biotin terminated DNA

DNA immobilisation solution prepared by adding 2 M sodium chloride, 1 M tris-HCl, 2 μM single stranded 5' biotin terminated DNA to nanopure H_2O. 10 μL of the DNA solution was drop cast onto the surface and left for 2 hours. Surfaces were washed with 10 mM tris-HCl (pH 7.6) three times.

4.4.7 DNA Hybridisation

Complementary DNA target solution (8 μL, 200 nM) in 0.5x PBS buffer was drop cast onto the glass surfaces for 15 minutes in a humid environment. Chips were washed with 0.5x PBS buffer for 30 minutes and nanopure H_20 for 30 minutes then dried with nitrogen.

4.4.8 Enzymatic Extension Method

0.2 μM DNA duplex (solution extension only), 200 nM polymerase, 10x reaction buffer (**Table 4.3**), 2mM magnesium sulfate ($MgSO_4$) and 0.5 mM deoxynucleotide triphosphates (dNTPs – dCTP, dATP, dTTP and dGTP) were mixed thoroughly and made up to the desired volume with nanopure water.

Thermocycling was performed using a VWR thermocycler by the method detailed below:

	30 seconds at 95 °C
N (number of cycles) x	30 seconds at 55 °C
	120 seconds at 72 °C

The products were cooled to 4 °C after the reaction and held at this temperature until removed from the thermocycler.

For solution extended DNA the products were purified using either a QIAGEN QIAquick PCR purification kit (QIAGEN, Manchester) or a Monarch PCR and DNA clean up kit (NewEngland Bio Labs Inc.). Manufacturer's protocol was followed.

Table 4.3. Polymerase Reaction Buffer

Buffer	Components
10x Pfu-Pol Z3 exo $^-$	200 mM Tris-HCl (pH 8.8 at 25 °C), 100 mM ammonium sulfate ($(NH_4)_2SO_4$), 100 mM postassium chloride (KCl), 1 % Triton X-100, 1 mg/mL bovine serum albumin (BSA) and 20 mM magnesium sulfate ($MgSO_4$)
1x ThermoPol	20 mM Tris-HCl (Ph 8.8 at 25 °C), 10 mM ammonium sulfate ($(NH_4)_2SO_4$), 10 mM potassium chloride (KCl), 2mM magnesium sulfate ($MgSO_4$) and 0.1 % Triton X-100

4.4.9 Fluorescence Detection

Hokkaido University

Surfaces were treated with 200x PicoGreen dye solution (8 μL) and left for 20 minutes at room temperature, covered with tin foil. The surfaces were washed with 3 mL nanopure H_2O and dried with nitrogen.

Surfaces were placed on top of a microscope slide and inserted into the imaging tray on a BioRad molecular imager Pharos FX. Fluorescence was set to the PicoGreen dye detection (488 nm) and measured as a low intensity sample.

Surfaces were treated with 200x PicoGreen dye solution (8 μL) and left for 20 minutes at room temperature, covered with tin foil. The surfaces were washed with 3 mL nanopure H_2O and dried with nitrogen.

Surfaces were placed on top of a microscope slide on an Axioshop 2 plus (Zeiss, Germany) image platform, set to filter 44, with a Plan-NEOFLUAR 10 x/ 0.3 objective lens (Zeiss). The surface was excited at 490 nm from a ebq100 mercury lamp (LEJ, Germany) and images were taken using an AxioCam HRm (Zeiss).

4.4.10 Contact Angle

Contact angle measurements were performed on a KSV Cam 101 (KSV Instruments Ltd., Finland) using built-in CAM 2008 software. A 2 μL drop of nanopure H_2O was dropped onto the surface. The software was used to estimate the angle of the water droplet on the surface. Ten measurements for each sample were recorded and the average calculated. For angles where there was more than a 2° difference in the left and right angle of the droplet, the results were discarded.

4.4.11 DNA Dehybridisation

extDNA surfaces were placed in 1 mL nanopure H_2O at 95 °C for 10 minutes. Surfaces were removed and transferred to a fresh 1 mL nanopure H_2O at 95 °C for 10 minutes. On removal surfaces were rinsed with 1 mL nanopure H_2O and dried with nitrogen. Time optimised to 40 minutes.

4.4.12 DNA Rehybridisation

Single stranded extDNA surfaces were immersed in 400 μL 100 μM target oligomer at 75 °C for 10 minutes and cooled slowly to room temperature. Surfaces were rinsed with 1 mL nanopure H_2O and dried with nitrogen. Temperature optimised to 45 °C.

4.4.13 Surface Photocleavage

The Lightningcure LC8, Hamamatsu Photonics) 365 nm UV source was set to 15 mW power. The photocleavable DNA surfaces were placed face down in 20 μL of nanopure H_2O and the UV source switched on. The samples were irradiated with UV for varying lengths of time and the fluorescence measured after each time point.

4.4.14 UV-Vis Spectroscopy

UV-Vis spectroscopy was carried out using a Thermo Scientific NanodropONE. 1.5 µL of sample was placed onto the stage and analysed between 1100 and 190 nm.

4.4.15 Agarose Gel Electrophoresis

DNA extension products were analysed by gel electrophoresis in TBE (tris, boric acid and EDTA) buffer. 1 % agarose (Melford Biolaboratories Ltd., Ipswich, UK) (1 g) was added to 1x TBE buffer (100 mL) and heated to dissolve. DNA samples were supplemented with 6x loading dye (2.5 % Ficoll-400, 11 mM EDTA, 3.3 mM Tris-HCl (pH 8 at 25 °C), 0.017 % sodium dodecyl sulfate (SDS) and 0.015 % bromophenol blue) (Thermo Scientific, UK) and GeneRuler 1 kilo base plus DNA ladder was used (Thermo Scientific, UK). The gels were run at 100 V, 100 mA, 10 W, for approximately 1 hour and post stained with 1 µg/µL ethidium bromide solution for 30 minutes. The gel was visualised using an ultra-violet transilluminator and imaged using UVIproMW software, version 11.02.

4.5 References

1. Rommens, J. *et al.* Identification of the cystic fibrosis gene: chromosome walking and jumping. *Science* **245**, 1059–1065 (1989).
2. Leach, F. S. *et al.* Mutations of a mute Homolog in Hereditary Nonpolyposis Colorectal Cancer. *Cell* **75**, 1215–1225 (1993).
3. Lunn, S. Exploiting DNA Surfaces For Sensing and Nanomaterial Applications. (Newcastle University, 2018).
4. Bracha, D., Karzbrun, E., Shemer, G., Pincus, P. A. & Bar-Ziv, R. H. Entropy-driven collective interactions in DNA brushes on a biochip. *Proc. Natl. Acad. Sci.* **110**, 4534-4538 (2013).
5. Sun, L., Svedhem, S. & Åkerman, B. Construction and modeling of concatemeric DNA multilayers on a planar surface as monitored by QCM-D and SPR. *Langmuir* **30**, 8432-8441 (2014).
6. Hatch, A., Sano, T., Misasi, J. & Smith, C. L. *Rolling circle amplification of DNA immobilized on solid surfaces and its application to multiplex mutation detection*. 35-40 (1999).
7. Wang, C. *et al.* DNA microarray fabricated on poly(acrylic acid) brushes-coated porous silicon by in situ rolling circle amplification. *Analyst* **137**, 4539–4545 (2012).
8. Brittain, W. J. & Minko, S. A structural definition of polymer brushes. *J. Polym. Sci. Part Polym. Chem.* **45**, 3505–3512 (2007).
9. Kobayashi, M. *et al.* Wettability and Antifouling Behavior on the Surfaces of Superhydrophilic Polymer Brushes. *Langmuir* **28**, 7212–7222 (2012).
10. Kreer, T. Polymer-brush lubrication: A review of recent theoretical advances. *Soft Matter* **12**, 3479–3501 (2016).
11. Mitomo, H. *et al.* Preparation and Characterization of Double-Stranded DNA Brushes via Surface-Initiated Enzymatic Polymerization. *J. Nanosci. Nanotechnol.* **17**, 8995-9001 (2017).
12. Nakamura, S. *et al.* DNA Brush-Directed Vertical Alignment of Extensive Gold Nanorod Arrays with Controlled Density. *ACS Omega* **2**, 46 (2017).
13. Mehne, J. *et al.* Characterisation of morphology of self-assembled PEG monolayers: a comparison of mixed and pure coatings optimised for biosensor applications. *Anal. Bioanal. Chem.* **391**, 1783–1791 (2008).
14. Glass, N. R., Tjeung, R., Chan, P., Yeo, L. Y. & Friend, J. R. Organosilane deposition for microfluidic applications. *Biomicrofluidics* **5**, 36501–365017 (2011).

15. Li, Y.-H. & Buriak, J. M. Dehydrogenative Silane Coupling on Silicon Surfaces via Early Transition Metal Catalysis. *Inorg. Chem.* **45**, 1096–1102 (2006).
16. Singer, V. L., Jones, L. J., Yue, S. T. & Haugland, R. P. Characterization of PicoGreen Reagent and Development of a Fluorescence-Based Solution Assay for Double-Stranded DNA Quantitation. *Anal. Biochem.* **249**, 228–238 (1997).
17. Whitfield, C. J., Turley, A. T., Tuite, E. M., Connolly, B. A. & Pike, A. R. Enzymatic Method for the Synthesis of Long DNA Sequences with Multiple Repeat Units. *Angew. Chem. Int. Ed.* **54**, 8971–8974 (2015).
18. Fishel, R. *et al.* The human mutator gene homolog MSH2 and its association with hereditary nonpolyposis colon cancer. *Cell* **75**, 1027–1038 (1993).
19. Bronner, C. E. *et al.* Mutation in the DNA mismatch repair gene homologue hMLH 1 is associated with hereditary non-polyposis colon cancer. *Nature* **368**, 258–261 (1994).
20. Papadopoulos, N. *et al.* Mutation of a mutL homolog in hereditary colon cancer. *Science* **263**, 1625–1629 (1994).
21. Zhou, X. P. *et al.* Determination of the replication error phenotype in human tumors without the requirement for matching normal DNA by analysis of mononucleotide repeat microsatellites. *Genes. Chromosomes Cancer* **21**, 101–7 (1998).
22. Schlotterer, C. & Harr, B. Microsatellite Instability. in *Encyclopedia of Life Sciences* (John Wiley & Sons, Ltd, 2004). doi:10.1038/npg.els.0000840
23. Berg, K. D. *et al.* Detection of microsatellite instability by fluorescence multiplex polymerase chain reaction. *J. Mol. Diagn. JMD* **2**, 20–28 (2000).
24. Losso, G., da Silveira Moraes, R., C Gentili, A. & Reason, I. Microsatellite instability–MSI markers (BAT26, BAT25, D2S123, D5S346, D17S250) in rectal cancer. *Arq. Bras. Cir. Dig. ABCD Braz. Arch. Dig. Surg.* **25**, 240–244 (2012).
25. Murphy, K. M. *et al.* Comparison of the Microsatellite Instability Analysis System and the Bethesda Panel for the Determination of Microsatellite Instability in Colorectal Cancers. *J. Mol. Diagn.* **8**, (2006).
26. Umar, A. *et al.* Revised Bethesda Guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability. *J. Natl. Cancer Inst.* **96**, 261-268 (2004).
27. Peterlinz, K. A., Georgiadis, R. M., Herne, T. M. & Tarlov, M. J. Observation of hybridization and dehybridization of thiol-tethered DNA using two-color surface plasmon resonance spectroscopy. *J. Am. Chem. Soc.* **119**, 3401–3402 (1997).

28. Gadsby, D. C., Vergani, P. & Csanády, L. The ABC protein turned chloride channel whose failure causes cystic fibrosis. *Nature* **440**, 477–483 (2006).
29. Tsui, L. *et al.* Cystic fibrosis locus defined by a genetically linked polymorphic DNA marker. *Science* **230**, 1054–1057 (1985).
30. Sheppard, D. N. & Welsh, M. J. Structure and Function of the CFTR Chloride Channel. *Physiol. Rev.* **79**, 23–45 (1999).
31. Childers, M., Eckel, G., Himmel, A. & Caldwell, J. A new model of cystic fibrosis pathology: Lack of transport of glutathione and its thiocyanate conjugates. *Med. Hypotheses* **68**, 101–12 (2007).
32. Riordan, J. R. *et al. Identification of the Cystic Fibrosis Gene: Cloning and Characterization of Complementary DNA*. *Source: Science, New Series* **245**, (1989).
33. Bartoszewski, R. A. *et al.* A synonymous single nucleotide polymorphism in DeltaF508 CFTR alters the secondary structure of the mRNA and the expression of the mutant protein. *J. Biol. Chem.* **285**, 28741–28748 (2010).
34. Huang, Y. M., Uppalapati, M., Hancock, W. O. & Jackson, T. N. Neutravidin micropatterning by deep UV irradiation. *Lab. Chip* **8**, 1745–1747 (2008).
35. Agasti, S. S., Liong, M., Peterson, V. M., Lee, H. & Weissleder, R. Photocleavable DNA Barcode−Antibody Conjugates Allow Sensitive and Multiplexed Protein Analysis in Single Cells Scheme 1. Schematic Illustration of the Light-Mediated Cellular Barcoding Strategy a. *J Am Chem Soc* **134**, 35 (2012).
36. Photo-Cleavable Modifiers and Their use with Oligonucleotides – LGC LINK Help Centre. Available at: https://helpcentre.linktech.co.uk/hc/en-us/articles/201707459-Photo-Cleavable-Modifiers-and-Their-use-with-Oligonucleotides. (Accessed: 18th January 2019)

Outlook

Outlook

To conclude, several recommendations that could lead to an increased understanding of the key aspects discussed in the previous chapters are outlined in the following sections.

Amplification

Chapter 2 reported the production of specific length designer DNA by exploiting the use of a Lonza™ double tier electrophoresis gels. The extraction of the DNA from these gel cassettes produced small volumes of user specified DNA lengths, however the concentration was often < 15 ng/μL, which makes further manipulation difficult. Amplification by qPCR was used on size recovered aliquots of DNA to try and achieve target concentrations of 50 μM in 100 μL. Unfortunately, the qPCR amplification proved difficult despite utilising various ligation techniques and altering different conditions to try and improve the amplification yield. Some conditions showed promising amplification results, in particular using a shortened annealing step, however several other parameters should be investigated to remove uncertain results.

The design of a suitable primer also requires further thought to facilitate the blunt end ligation technique in the replication and amplification of a repeat sequence target DNA. Further investigation into the qPCR amplification conditions is desired in order to optimise the heat-cool cycles. The conditions analysed throughout Chapter 2 (**Table 2.13**) should be a good guideline to aid in these investigations into the optimal amplification conditions.

DNA Designer Materials

A modified Whitfield enzymatic extension method successfully incorporated a new DNA section into a designed DNA sequence to generate ABCD-DNA. The synthesis of ABCD-DNA was performed using various sequences. Some sequences were difficult to extend due to their repetitive complex design, but the method could be tuned to accommodate this. To aid the extension of difficult sequences and ensure DNA growth, isothermal extension was incorporated into the method which extended the original oligoseed to a more stable structure prior to extension. The ABCD-DNA method was also utilised for the successful incorporation of phosphorothioate bases which created preferential spatial modification sites that were used for AuNP binding.

Future work to exploit this method could include increasing the length of the phosphorothioate modified ABCD-DNA by optimising the isothermal and Whitfield enzymatic extension methods. Also, a further step to improve the molecular combing of the DNA on the mica surface to ensure the DNA is aligned in a more linear fashion should be investigated, in order to ensure the phosphorothioate sites are more accessible for binding AuNPs. Finally, optimisation of the AuNP incubation method should be considered in order to understand the best method for AuNP deposition.

DNA Surfaces for Sensing

The extension of an immobilised oligoseed on a glass surface using the Whitfield enzymatic extension method successfully produced a DNA brush surface with enhanced fluorescence. $[GATC]_5$, BAT25 and CFTR sequences all extended off the surface to increase the fluorescence read-out compared to that of dsDNA. Furthermore, the BAT25 sequence showed that the DNA brush surface was able to determine between different base mutations from fluorescence measurements, however full reproducibility of the system was not achieved.

While the final results were able to determine between the different base mismatches, the dehybridisation procedure should still be improved by performing the process in buffer rather than nanopure water. It has been suggested by Irving *et al.*[1] that TE buffer afforded improved dehybridisation conditions. Furthermore, the rehybridisation procedure should also be examined further by using PBS buffer to achieve optimal target DNA annealing.

General Outlook

If the potential of DNA as a nanomaterial is to be fully realised, methods to synthesise large volumes of designer DNA are essential. Although, DNA can be commercially produced by automated synthesis, it is limited in length and sequence. Despite the relatively low cost of oligo synthesis, € 140 for 0.2 μM duplex oligo (Eurofins, February 2019), further reductions are necessary. Enzymatic methods have been exploited to synthesise DNA, however they are limited by the commercial availability of DNA templates and are restricted by highly repetitive sequences or the incorporation of modifications. Touchlight Ltd. have started to generate DNA as a material through their Doggybone DNA (dbDNA™) method,[2,3] yet this method can only form DNA concatemers or long single stranded DNA, not long dsDNA. Therefore, the Whitfield enzymatic

extension method, coupled with size recovery and amplification, is an advantageous method to enable cost-effective synthesis of user-designed DNA. Furthermore, alongside its use as a material, DNA also has the potential for use in ink formulation.[4] For ink generation and printing, large volumes of DNA are required to process DNA on an accessible scale. Therefore, in order to advance DNA based technologies, designer DNA must make the transition from a fine chemical to a commodity on large scale production.

It is also essential to continually advance the types of DNA available in order to generate new structures, materials and to incorporate new functionalities. The types of enzymatically synthesised DNA, reported here, could be used in a range of self-assembly approaches for nanomaterial fabrication within one DNA strand. Enzymatically modified designer ABCD-DNA has the potential to fabricate functionalised DNA origami for use in drug binding and release studies.

The fabrication of DNA microarray sensors, using the reported DNA brush platform, would improve the detection of genetic mutations allowing for quicker patient diagnosis and treatment. Microarray structures of the DNA surface could be achieved by inkjet printing of DNA probes at localised areas. The modification of an optical fibre with extended DNA is one possible application of this technology in the optical sensing of DNA. The combination of the enhanced signal from the extended DNA and the sensitivity of an optical fibre could be a promising route for more accurate and conclusive determination of diseases caused by single base mutations.

References

1. Irving, D., Gong, P. & Levicky, R. DNA Surface Hybridization: Comparison of Theory and Experiment. *J. Phys. Chem. B* **114**, 7631–7640 (2010).
2. Allen, A. *et al.* Linear doggybone DNA vaccine induces similar immunological responses to conventional plasmid DNA independently of immune recognition by TLR9 in a pre-clinical model. *Cancer Immunol. Immunother.* **67**, 627–638 (2018).
3. Patel, A. *et al.* Novel synthetic plasmid and Doggybone™ DNA vaccines induce neutralizing antibodies and provide protection from lethal influenza challenge in mice. L. *Hum. Vaccines Immunother.* **11**, 1972–1982 (2015).
4. Nurdillayeva, R. N. *et al.* Inkjet printing and electrical characterisation of DNA-templated cadmium sulphide nanowires. *Nanotechnology* **29**, 135704 (2018).