UNIVERSITY OF SOUTHAMPTON

FACULTY OF ENGINEERING, SCIENCE AND MATHEMATICS

School of Chemistry

Factors Affecting the Read-Out of DNA Suspension Arrays

by

Hugo Filipe Pedro Martins

Thesis for the degree of Doctor of Philosophy

March 2007

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF ENGINEERING, SCIENCE AND MATHEMATICS

SCHOOL OF CHEMISTRY

Doctor of Philosophy

FACTORS AFFECTING THE READ-OUT OF DNA SUSPENSION ARRAYS

by Hugo Filipe Pedro Martins

DNA suspension arrays provide the means for a high-throughput parallel analysis. However, hybridisation on solid supports has not been studied as extensively as in solution. The relationship between fluorescence read-out, hybridisation strength and selectivity, and solution and washing conditions of the array platform was investigated. Fluorescently labelled oligonucleotide targets were hybridised to oligonucleotide probes immobilised on avidin coated GMA beads. The DNA duplexes formed by the hybridisation were designed to contain a comprehensive set of single base mismatches centrally located in the double helix. The beads were subjected to a series of post-hybridisation washes with increasing stringency and duplex stability monitored by single bead fluorescence. The solution T_m for all duplexes was determined in a range of solvent conditions representing the different washing steps. It was found that the thermodynamic stability trend predicted by the solution T_ms did not correlate the fluorescence read-out exhibited by the immobilised duplexes. Results suggested that the hybridisation discrimination achieved between complementary and mismatched duplexes was due to faster kinetics of dissociation of the mismatches under the washing conditions employed rather than discrimination based in thermodynamic stability. Solid phase single base extension and restriction endonuclease digestion of bead immobilised DNA duplexes were also explored. The performance of DNA polymerases and an endonuclease was investigated and compared in solution and on solid phase. Results showed that nucleotide incorporation is more effectively achieved in solution than on solid phase while solution and solid phase restriction digestions yielded similar results.

Contents

Acknowledgments

I would like to thank Dr. Cameron Neylon for his supervision and guidance throughout the course of my PhD and his immense help in the process of writing this thesis. I would like to thank all those involved in the 4G project, which directly or indirectly contributed to the success of my work. I want also to express my gratitude to Professor Tom Brown and his research group for allowing me to use the spectrophotometer.

I thank all my colleagues from the wet chemistry group, past and present, Gabriel, Shannon, Joseph, Graham and Rohan for all their help and support and for making the lab a pleasant place to work in. I thank Rob for being such a nice person and for the help and suggestions in the writing of this thesis.

I would like to thank all my friends who made Southampton, outside of the School, an enjoyable place to be; Gabriel, Joseph, Lilyan, Nefeli and Alex. Thank you for all the good moments we spent together. We shared great fun, laughter, great meals and the occasional appendectomy (!!!). To you my friends, *thank you*, *gracias*, *ευχαριστώ* and 多謝 .

And finally, I would like to say *obrigado* to my mother Lisete, to my friends Anita, Paulo, Violeta and Filipa for their constant support, enormous patience over the course of these years and for pushing me forward. Thank you for always being on my side and for taking care of me!

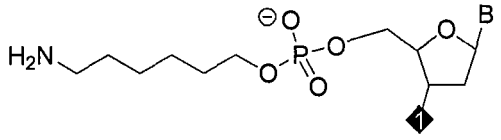
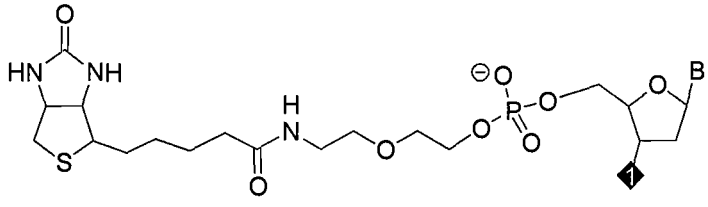
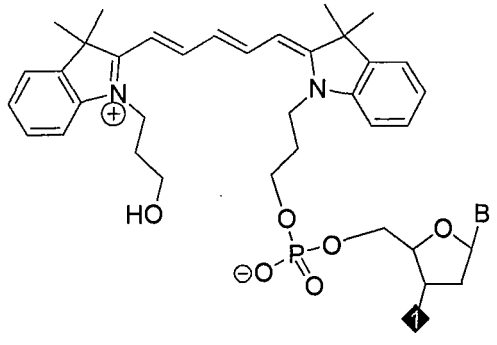
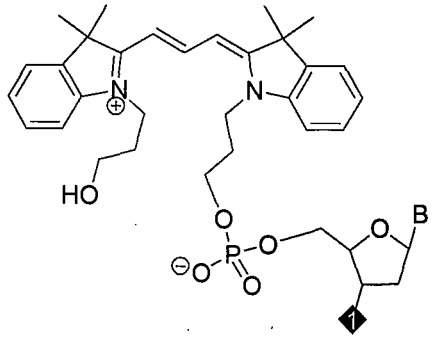
Cursum perficio

Abbreviations

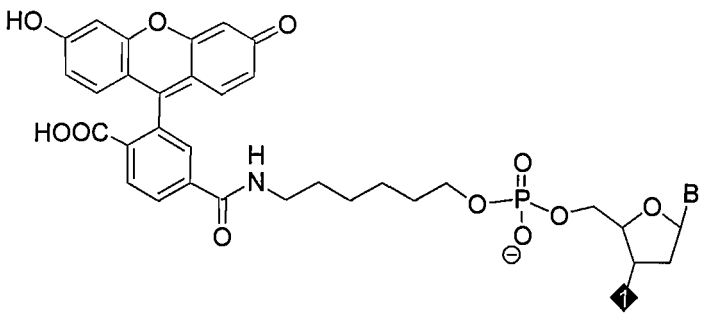
530/30 FITC-A	Band-pass filter used in FACS for fluorescein detection
575/26 PE-A	Band-pass filter used in FACS for Cy3 detection
660/20 APC-A	Band-pass filter used in FACS for Cy5 detection
695/40 PerCP-Cy5-5-A	Band-pass filter used in FACS for FRET detection using FRET pair, Fluorescein and Cy5
1xSSC	0.15 M NaCl, 15 mM Na_3citrate, pH 7.0
2xSSC	0.3 M NaCl, 30 mM Na_3citrate, pH 7.0
5xSSPE	0.75 M NaCl, 50 mM NaH_2PO_4, 5 mM EDTA, pH 7.4
A	Adenine
Abs	Absorbance
Aoct	Amino-octanoic acid
ATP	Adenosine triphosphate
C	Cytosine
cDNA	Complementary DNA
Count	Number of recorded events (particles detected) by FACS
Cy	Cyanine (dye)
Cy5-dUTP	Cy5 - 5 - NH_2 – propargyl – 2' deoxyuridine 5'- triphosphate
ddATP	2', 3'- dideoxyadenosine 5'- triphosphate
ddCTP	2', 3'- dideoxycytidine 5'- triphosphate
ddGTP	2', 3'- dideoxyguanosine 5'- triphosphate
ddNTP	2', 3'- dideoxynucleotide 5'- triphosphate
ddTTP	2', 3'- dideoxythymidine 5'- triphosphate
DIC	N,N'- Diisopropylcarbodiimide
DMF	N,N - dimethylformamide
DNA	Deoxyribonucleic acid
dNTPs	2'- deoxynucleotide 5'- triphosphate
EDC	1-ethyl-3-(3-dimethylaminopropyl) carbodiimide
EDTA	Ethylenediaminetetraacetic acid
EtOH	Ethanol
FACS	Flow activated cell sorter/sorting
FITC 530/30-A	See 530/30 FITC-A

Fmoc	9-Fluorenylmethoxycarbonyl
FRET	Fluorescence resonance energy transfer
FSC-A	Forward scatter in FACS analysis – provides information about particle size
G	Guanine
GMA	Glycidyl methacrylate
H	Hydrogen
HeNe	Helium-Neon
HOBt	Hydroxybenzotriazole
HPLC	High pressure liquid chromatography
K_D	Equilibrium dissociation constant
LDPS	Laser diffraction particle size
MeCN	Acetonitrile
MeOH	Methanol
mRNA	Messenger RNA
NMR	Nuclear magnetic resonance
OLA	Oligonucleotide ligation assay
PAGE	Polyacrylamide gel electrophoresis
PBS	10 mM K_2HPO_4, 1.8 mM NaH_2PO_4, 0.15 M NaCl, pH 7.4
PCR	Polymerase chain reaction
Pfu Turbo	*Pyrococcus furiosus*
PEG	Polyethylene glycol
PMT	Photo-multiplier tube
PNA	Peptide nucleic acid
PPi	Pyrophosphate
RNA	Ribonucleic acid
RFLP	Restriction fragment length polymorphism
Rox	Carboxyrhodamine
RSP	Restriction site polymorphism
SBCE	Single base chain extension
SDS	Sodium dodecyl sulphate
SE	Standard Error
SEM	Scanning electron microscopy

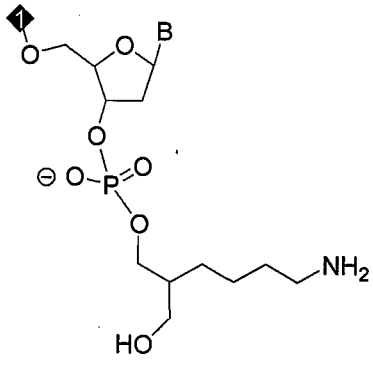
SNP	Single nucleotide polymorphism
SSC 488/10-A (SSC-A)	Side scatter in FACS analysis – provides information about particle morphology
SU8	Negative photoresist, originally developed and patented by IBM
T	Thymine
TAMRA	Carboxytetramethylrhodamine
Taq	*Thermus Aquaticus*
TBE	Tris-borate-ethylenediamine tetracetic acid
TEA	Triethylamine
TEMED	N,N,N',N'-Tetramethylethylenediamine
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
T_m	Melting temperature
Tris	Tris(hydroxymethyl)aminomethane
U	Uracil / units of enzyme
UV	Ultraviolet
wt	Wild type

Oligonucleotide modification	Structure (1 refers to oligonucleotide chain)
5' NH_2	
5' Biotin	
5' Cy5	
5' Cy3	

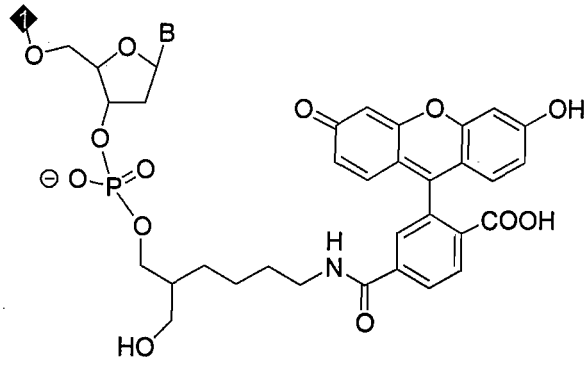
5' Fluorescein



3' NH_2



3' Fluorescein



1. Introduction

1.1 Composition and structure of nucleic acids

DNA (deoxyribonucleic acid) and RNA (ribonucleic acid) are condensed polymers of monomeric units called nucleotides. Figure 1.1 outlines the formation of the building blocks of DNA. The union of a nitrogenous base to a pentose sugar (2-deoxyribose or ribose, for DNA or RNA respectively) forms the nucleoside, and the further association to a phosphate group forms the nucleotide. The nitrogenous bases found in nucleic acids are either monocyclic pyrimidines, cytosine and thymine (or uracil, found in RNA), or bicyclic purines, adenine or guanine (Figure 1.2).

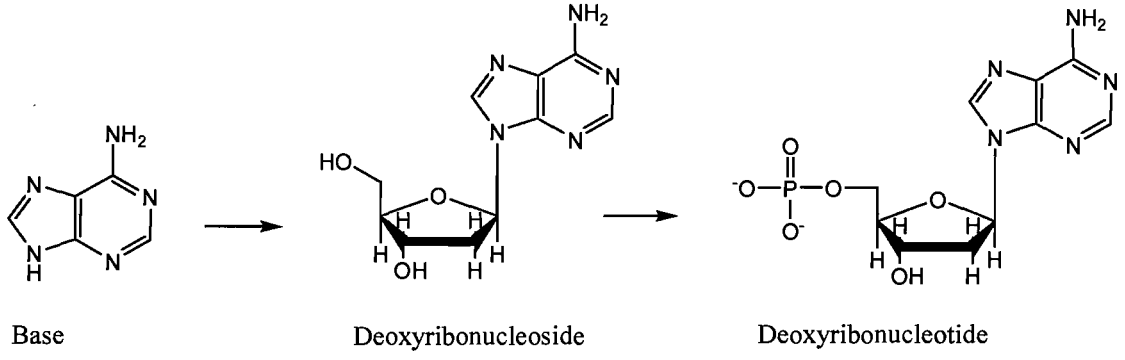


Figure 1.1 Structural units found in nucleic acids. A base (Adenine as example) in conjunction with a sugar forms a nucleoside and the addition of the phosphate group makes the nucleotide.

Early work by Chargaff[1] on the comparison of an extensive variety of DNA sources, led him to conclude that although the base composition could vary within the duplex, the amount of adenine was the same as the amount of thymine, and the amount of guanine was the same as the amount of cytosine. Watson and Crick[2, 3] observed that base pairing was made through the association of one purine to one pyrimidine: Guanine would pair with cytosine, and adenine with thymine (Figure 1.3).

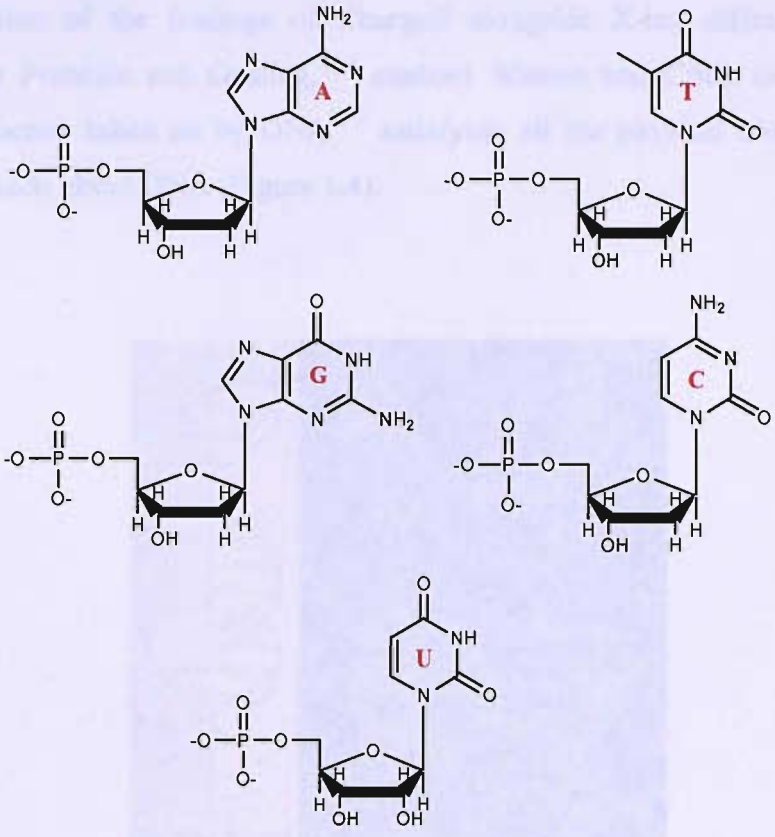


Figure 1.2 Structures of nucleotides found in DNA and RNA; Adenine (A), Thymine (T), Cytosine (C), Guanine (G) and Uracil (U).

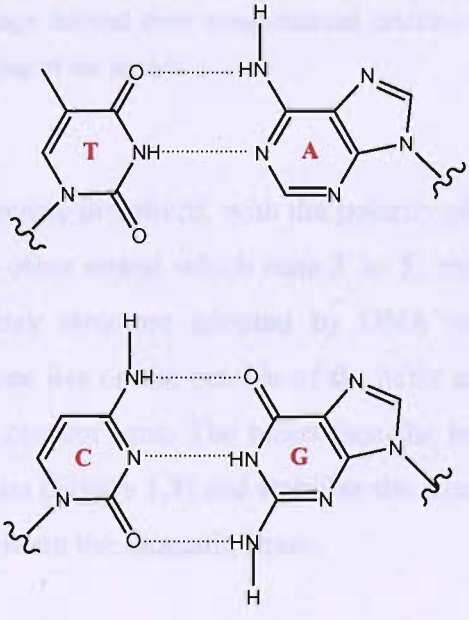


Figure 1.3 Watson and Crick base pairing found in DNA. The specificity of the base pairs relies on the number of hydrogen bonds between bases.

The combination of the findings of Chargaff alongside X-ray diffraction patterns undertaken by Franklin and Gosling,[4, 5] enabled Watson and Crick to postulate the secondary structure taken on by DNA[2, 3] satisfying all the physical observations and implications made about DNA (Figure 1.4).

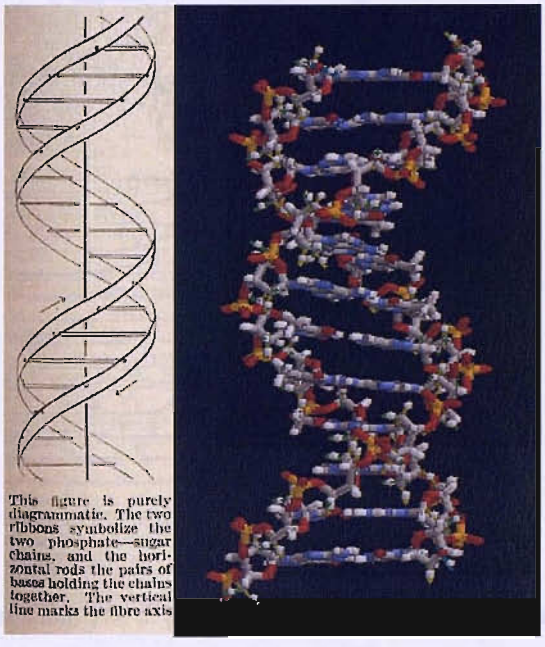


Figure 1.4 DNA helicoidal structure. On the left the diagrammatic structure presented by Watson and Crick[3] and on the right an image derived from computational calculations,[6] displaying the atoms and bonds involved in the intertwining of the strands.

The two strands run in opposite directions, with the polarity of one chain running 5' to 3' (Figure 1.5) paired to the other strand which runs 3' to 5', meaning the double helix is anti-parallel. The secondary structure adopted by DNA means that the negatively charged phosphate backbone lies on the outside of the helix and can be readily solvated by water or stabilised by counter ions. The bases face the inside of the helix forming Watson and Crick base pairs (Figure 1.3) and stabilise the construct by hydrophobic and π stacking interactions between the aromatic bases.

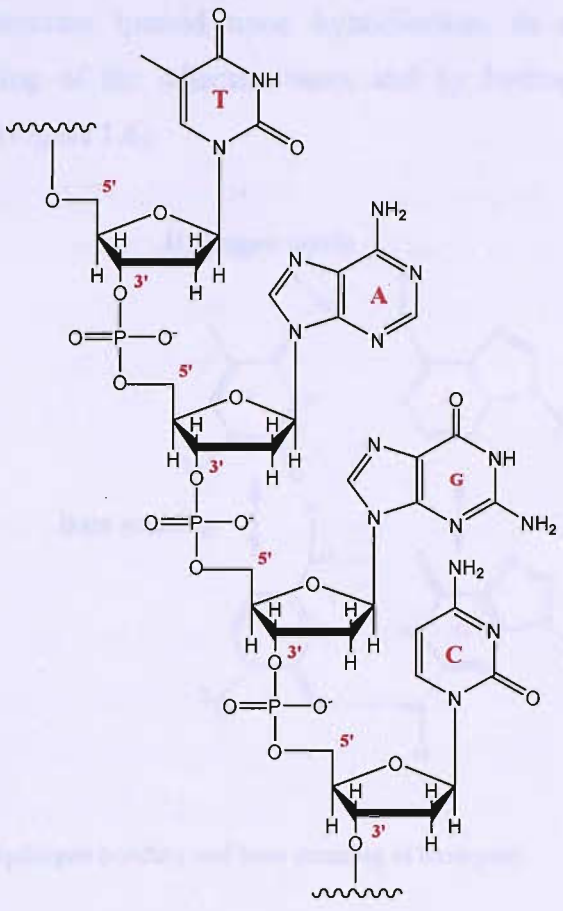


Figure 1.5 5' to 3' chain elongation. The 5' phosphate of one nucleotide links to the 3' OH of the pentose of the next nucleotide. DNA is synthesised from the 5' to the 3' direction.

1.1.1 The biophysics and thermodynamics of hybridisation

Hybridisation is the formation of double stranded nucleic acids. It is a multi-step process that is believed to begin with nucleation, that is, the initial pairing of a small number of bases in a transient intermediate. If the two strands are complementary, duplex formation propagates outward from the nucleation point in a process known as zippering.[7] Although zippering is reversible, the probability of strand separation diminishes as the double stranded region grows. The stability of the duplex construct determines its dissociation rate.

The DNA duplex structure formed upon hybridisation, as already mentioned is stabilised by π stacking of the adjacent bases and by hydrogen bonding between complementary bases (Figure 1.6).

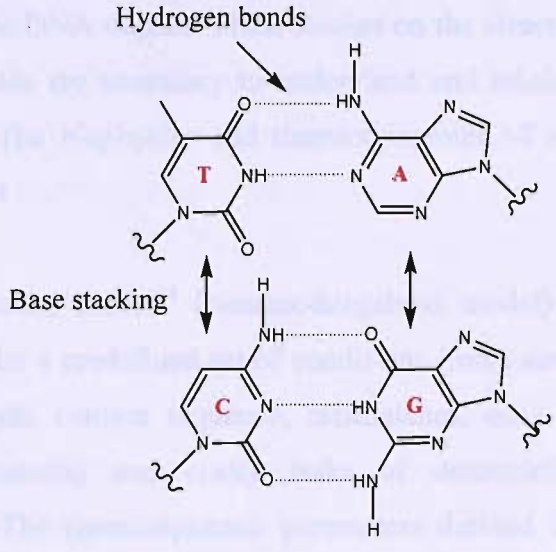


Figure 1.6 Schematics of hydrogen bonding and base stacking in base-pairs.

Figure 1.6 shows that G:C base pairs form three hydrogen bonds in contrast to T:A base pairs which form only two. This means that higher G:C content increases the melting temperature (T_m) and stability of the duplex formed.[8] T_m is defined as the temperature at which half of the double stranded DNA dissociates into single stranded DNA.[9] The thermodynamic stability of DNA also depends on the length[10] and sequence order or context. Two sequences with the same G:C content can have different stabilities.

The influence of metal ions has been shown to impart dramatic influence in the T_m of DNA.[11] A DNA duplex is a highly negatively charged polyelectrolyte due to the charges located on the phosphate backbone. This leads to strong repulsive electrostatic interactions between the phosphate backbones. On the other hand, the presence of counter ions significantly stabilises the DNA duplex.[12] This electrostatic interaction is more effective with monovalent cations. However, high concentrations of monovalent salt have a destabilising effect. The salt anions lead to the denaturation of DNA, acting as hydrophobic bond-breaking agents when they are present at high concentrations. Divalent cations have different effects on the stabilisation of DNA, for example, Mg^{2+}

increases T_m while Cu^{2+} has the opposite effect.[13] This is probably explained by interaction of the divalent cations with different sections of the duplex.

The presence, identity, number and location of mismatched base pairs, can also affect strongly the stability of the DNA duplex. Thus, studies on the structure and stability of mismatches in nucleic acids are necessary to understand and minimise the chance of incorrect base pairing.[14] The biophysics and thermodynamics of mismatches will be further studied in chapter 4.

A widely used computational model[14] (Nearest-neighbour model) for predicting the stability of DNA (T_m) under a predefined set of conditions (ionic strength, temperature, concentration, probe length, context sequence, mismatches, etc.) was developed, to overcome the time-consuming and costly tasks of determining T_m under all experimental conditions. The thermodynamic parameters derived by the model have provided a major impact in DNA technology.[15]

1.2 The importance of sequencing nucleic acids

Sequencing is the process of determining the identity and order of nucleotides in a given DNA or RNA fragment. The demand for this sequence information comes from the understanding that DNA encodes most, if not all, of the information required to make an organism work. There is an intimate relation between the genetic information contained in DNA and the functions of the cell. This genetic information is first transcribed to the intermediate messenger RNA (mRNA) and then translated into proteins which mediate all cellular processes.[9]

Errors in DNA transcription such as deletions and mutations will lead to genes being differently expressed. These errors will encode different proteins which might have an impact on disease phenotype[16] leading to genetic disorders such as cystic fibrosis,[17] sickle cell anaemia[18, 19] and cancer for example.[9, 20]

The most common sources of genetic variation between humans are single nucleotide polymorphisms (SNPs). SNPs are DNA sequence variations that occur when a single nucleotide (A, T, C, or G) in the genome sequence is altered. SNPs reflect past mutations that were mostly (but not exclusively) unique events and two individuals sharing a variant allele are thereby marked with a common evolutionary legacy.[21] SNPs occur on average at a density of one SNP per 1.91 kilobases in human genomic DNA.[21-23] There are believed to be around three million[16, 21-24] SNPs in total. They may not all directly account for detectable phenotypes,[19] but because of their high frequency and widespread nature they can function as genomic markers for susceptibility and resistance genes in diseases with simple and complex genetic backgrounds.[16, 24]

Sequencing DNA is critical to understanding the hereditary nature of disease as well as evolution in general. By sequencing and reading the genome of an individual at a specific gene, that individual may be screened for mutations that may result in disease, allowing suitable treatments to be identified before symptoms manifest themselves. The ability to screen a genome rapidly has many applications aside from disease screening including disease carrier screening, individual/species identification, paternity testing and genotyping.[9, 16]

1.3 Polymerase chain reaction

DNA obtained from biological samples usually requires amplification of a target region in order to increase the concentration to detectable levels.[18] The polymerase chain reaction (PCR), developed by Mullis,[25, 26] is generally the amplification method of choice (Figure 1.7), as it provides a simple method for replicating many copies of a specific segment of DNA.[9] The solution mixture of target DNA to be replicated also comprises excess synthetic oligonucleotide primers, a heat stable DNA polymerase and deoxynucleotide triphosphate bases (dNTPs). Labelled dNTPs may also be employed in order to achieve fluorescent PCR products.

Two synthetic primers are used, each complementary to the different strands of the DNA duplex and binding at either end of the target region to be amplified. The reaction

mixture is initially heated to 95 °C to denature all hybridised DNA and then rapidly cooled to a lower temperature to promote the hybridisation of the primers and the targets. These temperatures are usually about 50 °C to 70 °C, depending on the T_m of the primer-target duplexes. The temperature is then increased up to the optimal temperature required for enzyme activity, and the polymerase elongates both primers in the 5' to the 3' direction, producing thus 2 copies of the target.

This cycle is repeated n times in order to produce 2^n target copies. PCR is a chain reaction because newly synthesised DNA strands act as templates for further DNA synthesis in subsequent cycles.[9] To ensure the primer-target hybridisation is favoured over the re-formation of the target duplexes, the concentration of the primers is present in large excess.

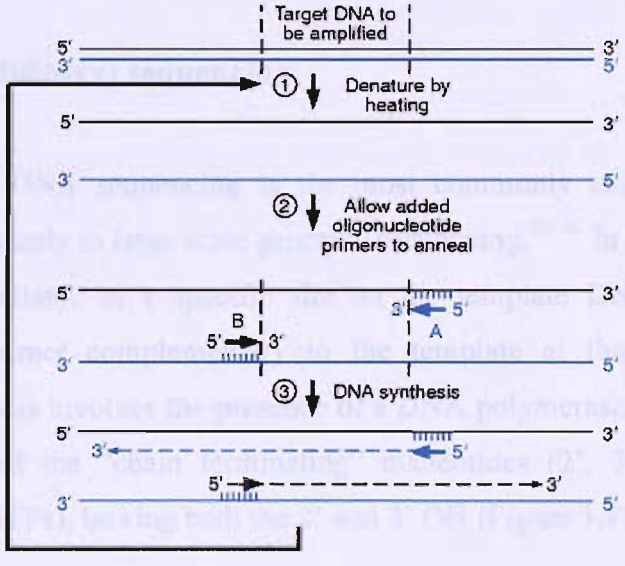


Figure 1.7 A schematic representation of PCR. Primers A and B are complementary to DNA sequences located on opposite DNA strands and flanking the region to be amplified. This cycle is usually repeated 25 times in an hour producing 2^{25} target copies. Adapted from ref. 9.

1.4 Modern technologies for genetic analysis

There are many technologies widely available for genetic analysis. Despite this large number of technologies, no technique has become a widely accepted standard for the identification of SNPs and detection of mutations.[27] These technologies for DNA analysis can be divided in two basic categories; Enzyme-mediated and direct hybridisation.[28] The first category refers to techniques of DNA analysis assisted by enzymatic activity. The second category relies on the sequence specific detection of DNA by direct hybridisation. The following sections describe some of the most commonly used technologies for DNA analysis representative of each category.

1.4.1 Enzyme-mediated methods for DNA analysis

1.4.1.1 Sanger (dideoxy) sequencing

Sanger (dideoxy) DNA sequencing is the most commonly used method for DNA sequencing, particularly in large scale genomic sequencing.[29, 30] In Sanger sequencing,[30, 31] extension is initiated at a specific site on the template DNA by using a short oligonucleotide primer complementary to the template at that region. The DNA amplification process involves the presence of a DNA polymerase, the deoxynucleotide bases (dNTPs) and the "chain terminating" nucleotides (2', 3'- dideoxynucleotide triphosphates, ddNTPs), lacking both the 2' and 3' OH (Figure 1.8).

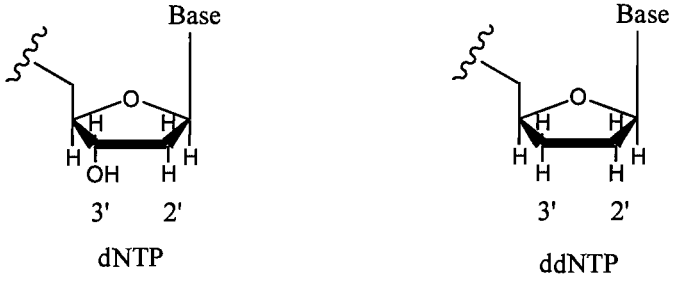


Figure 1.8 Comparison between the structures of dNTPs and ddNTPs. ddNTPs contain no 3' OH and the chain extension stops after their incorporation.

Limited incorporation of the chain terminating nucleotides by the DNA polymerase results in a series of related DNA fragments that are terminated only at positions where a particular ddNTP is used. These fragments differ in length, the smallest fragment being one nucleotide longer than the primer and the longest being the entire length of the sequence.[31] Comparison of the different sized fragments obtained in the four reactions, resolved by gel or capillary electrophoresis, enables the template sequencing.

There are two popular automated variations of Sanger sequencing;[29] dye-primer[32, 33] and dye-terminator[29, 34] sequencing. In dye-primer sequencing, the PCR primer is 5' fluorescently labelled. This method requires four separate extension reactions and four dye-labelled primers for each template. The fact that sequencing reactions must be carried in quadruplicate reduces throughput and increases labour costs. This method benefits from sequencing electropherograms with very even peak heights.[29] In dye-terminator sequencing, the four ddNTPs are added, each labelled with a different fluorophore (Figure 1.9). Elongation products are then separated by electrophoresis in a single gel-lane and the fluorescence wavelength emitted from each oligonucleotide indicates the 3' terminal nucleotide and sequence is evaluated (Figure 1.10). By using a different fluorescent label for each ddNTP the need for the reaction to be repeated in quadruplicate is avoided. This is most convenient and faster.

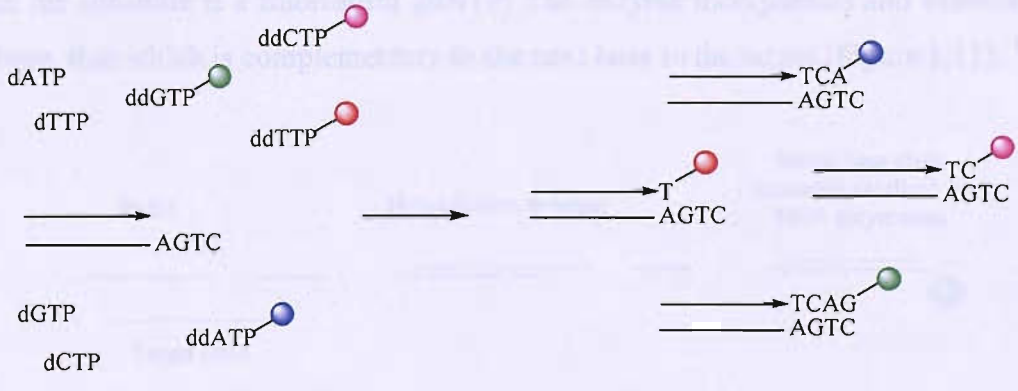


Figure 1.9 Dye-terminator Sanger Sequencing

The major disadvantage of the dye-terminator sequencing is the production of electropherograms with uneven peak heights, due to a template and enzyme dependent difference in the incorporation of the large dye chain-terminators. This problem has been significantly reduced with the introduction of new enzymes and dyes that minimize incorporation variability.[29]

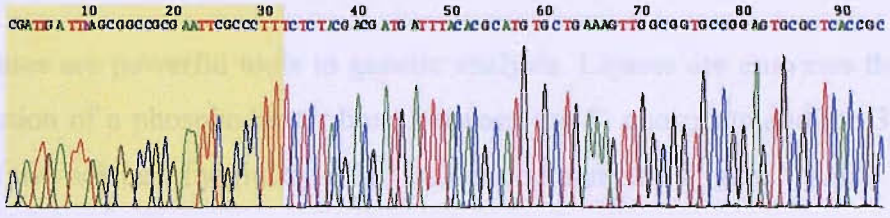


Figure 1.10 Dye-terminator electropherogram. The different four labels are detected and represented as 'peaks' of different colours. This determines the base sequence, shown at the top. Adapted from ref. 35.

1.4.1.2 Single base chain extension

Single base chain extension was developed based on the chain-terminator strategy. A DNA polymerase uses an oligonucleotide probe as a primer in an extension reaction in which the substrate is a fluorescent ddNTP. The enzyme incorporates and extends only one base, that which is complementary to the next base in the target (Figure 1.11). [16, 36]

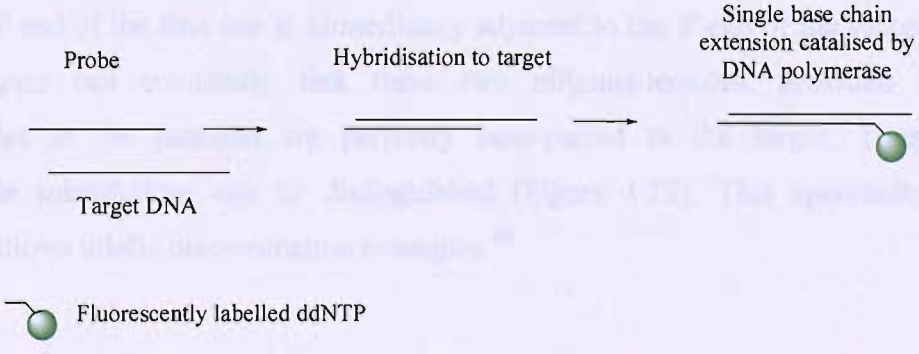


Figure 1.11 Detection of unlabelled DNA targets by single base chain extension.

This method is particularly useful for identifying SNPs[18, 37] and allows a multiplex approach. [28, 38] By choosing primers adjacent to polymorphic loci, extension in the presence of four differently labelled ddNTPs, reveals the identity of the complementary bases at the polymorphic sites.

1.4.1.3 Oligonucleotide Ligation Assay

DNA ligases are powerful tools in genetic analysis. Ligases are enzymes that catalyse the formation of a phosphodiester bond between the 5' phosphate and the 3' hydroxyl groups of two adjacent hybridised oligonucleotide fragments (Figure 1.12).

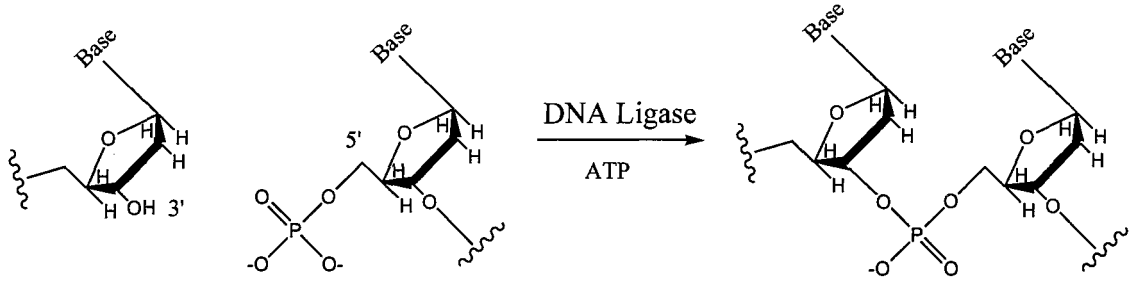


Figure 1.12 Formation of a phosphodiester bond between the 5' phosphate and the 3' hydroxyl groups by a DNA ligase.

The enzymatic activity of DNA ligases led to the development of the oligonucleotide ligation assay (OLA).[39] Two oligonucleotide probes are hybridised to target DNA, such that the 3' end of the first one is immediately adjacent to the 5' end of the second probe. DNA ligase can covalently link these two oligonucleotides, provided that the nucleotides at the junction are perfectly base-paired to the target. Thus, single nucleotide substitutions can be distinguished (Figure 1.13). This specificity of the ligation allows allelic discrimination strategies.[40]

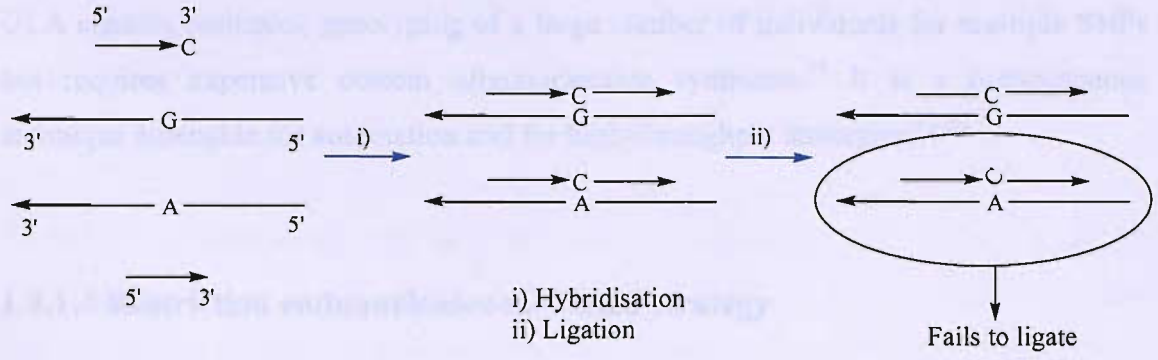


Figure 1.13 Schematics of a DNA ligase strategy.

The ligation process forms one new DNA molecule, distinct from any sequence previously present in the reaction and ready to be detected via DNA amplification procedures.[40, 41] A single-base mismatch prevents ligation and amplification and is also distinguished.

Fluorescence Resonance Energy Transfer (FRET) based ligation strategies have been reported (Figure 1.14).[42, 43] Both oligonucleotide probes are labelled with different dyes, one bearing a donor and the other an acceptor. Ligation forms a DNA molecule that can be excited at the donor's excitation wavelength, leading to emission by the acceptor via FRET. This allows the monitoring of the progress of the reaction by a change in fluorescence signal.

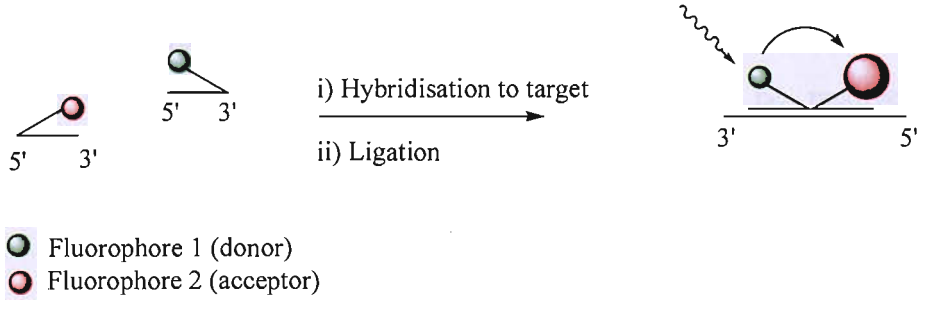


Figure 1.14 FRET-based ligation strategy.

OLA enables multiplex genotyping of a large number of individuals for multiple SNPs but requires expensive custom oligonucleotide syntheses.[16] It is a homogeneous technique amenable for automation and for high-throughput strategies.[40, 42]

1.4.1.4 Restriction endonucleases-mediated strategy

Restriction endonucleases are fundamental tools in genetic engineering. These enzymes cleave the DNA duplex at specific recognition nucleotide sequences (Figure 1.15). The position of these specific restriction sites are determined by the identity of the restriction endonuclease. The cleaved DNA fragments can be joined using a DNA ligase, which re-forms the phosphodiester bonds of DNA.

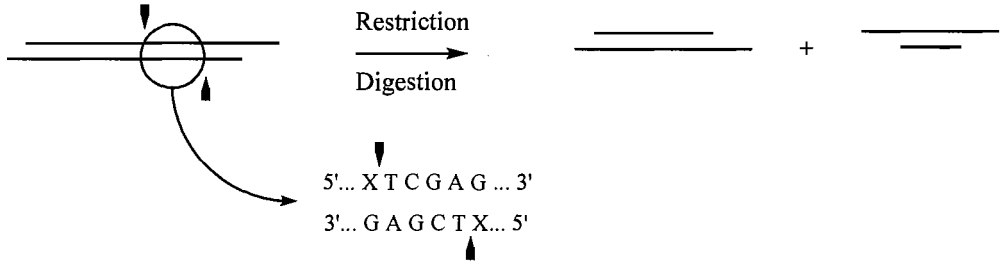


Figure 1.15 Schematics of a restriction endonuclease digestion.

A common use for restriction endonucleases is to generate a "fingerprint" of a particular DNA molecule. Because of the sequence specificity of restriction endonucleases, these enzymes can cut DNA into discrete fragments which can be resolved by gel electrophoresis. This pattern of DNA fragments generates a "DNA fingerprint," and each DNA molecule has its own fingerprint.

The specificity of this enzymatic digestion activity can be exploited for allelic discrimination.[44] Alleles will possess or lack a recognition site for a specific restriction endonuclease and will therefore display restriction site polymorphism (RSP). RSPs can be assayed by restriction digestion using an appropriate endonuclease. The presence or absence of a recognition site will lead to a difference in size of allelic restriction fragments. This is known as restriction fragment length polymorphisms (RFLPs).[9]

RFLPs are the basis of PCR-RFLP[45, 46] assays, which involve amplification of the target sequences prior to endonuclease digestion. It is a technique amenable for multiplexing.[47] However, PCR-RFLP is labour-intensive and time consuming.

Restriction endonucleases can also be used in conjunction with DNA ligases for detecting mutations[48] and for DNA sequencing.[49, 50] SNP analysis through single base chain extension of digested fragments has also been reported.[51, 52] After endonuclease digestion, the 5' overhang of the fragment is extended using a DNA polymerase.

1.4.1.5 Scorpions

A Scorpion primer consists of a probe sequence held in a hairpin loop conformation by a stem at its 5' and 3' termini. A fluorophore is attached to the 5' end of one arm of the stem, and a quencher to the 3' end of the other arm. The 3' terminus of the stem is attached to the 5' end of a PCR primer by a PCR stopper, which prevents copying of the probe by *Taq (Thermus Aquaticus)* DNA polymerase (Figure 1.16).[53]

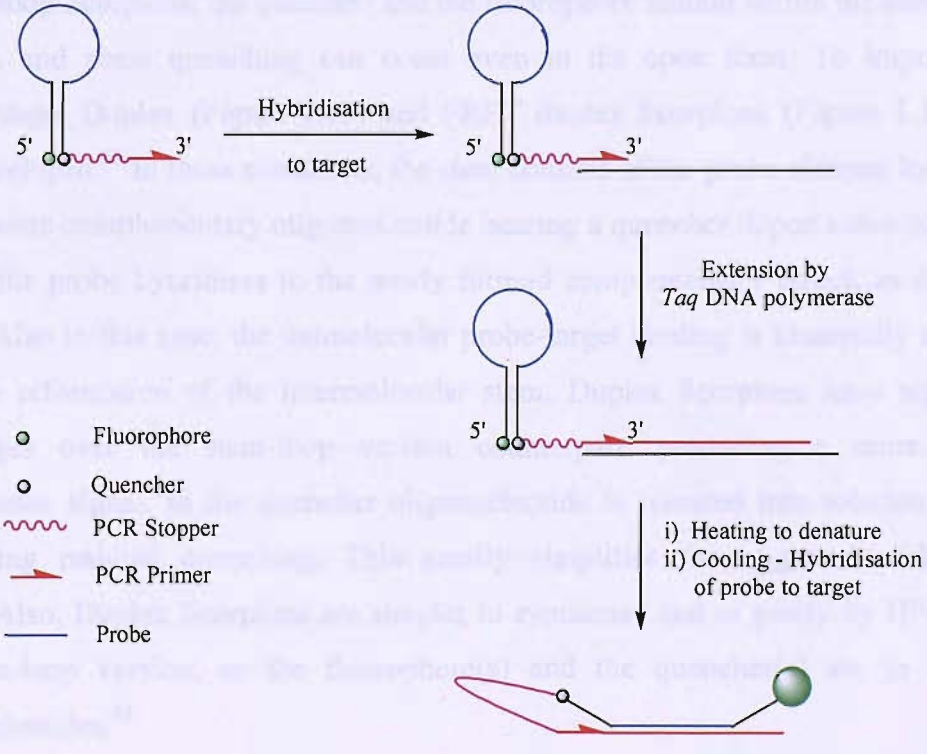


Figure 1.16 Schematics of a Scorpion detection system.

When the Scorpion primer is extended upon hybridisation to target, the result is a DNA strand, a section of which is complementary to the probe sequence. This region of complementarity is longer than that of the stem and its hybridisation is favoured over the reformation of the stem. This prompts a conformational reorganisation of the Scorpion; the hairpin loop opens, which forces the fluorophore and the quencher to move away from each other, thereby releasing fluorescence.[53, 54]

The Scorpion becomes an integral part of the amplified target. Since both probe and target are now in the same molecule, probe-target binding is kinetically favoured over duplex re-hybridisation and thermodynamically favoured over intrastrand secondary structures. This results in the Scorpions primers having advantages over other assays. The speed of this unimolecular probing event generates a rapid signalling technology and high signal to noise ratios which allow faster assays when compared to bimolecular processes.[53] Scorpions technology can be used in allelic discrimination and is effective in SNP genotyping.[55] Also, because Scorpions probe in a strand specific manner multiplexing of reactions is possible.[17, 53]

In stem-loop Scorpions, the quencher and the fluorophore remain within the same strand of DNA and some quenching can occur even in the open form. To improve this disadvantage, Duplex (Figure 1.17) and FRET duplex Scorpions (Figure 1.18) have been developed.[55] In these constructs, the stem consists of the probe element hybridised to a separate complementary oligonucleotide bearing a quencher. Upon extension of the primer, the probe hybridises to the newly formed complementary target, as described before. Also in this case, the unimolecular probe-target binding is kinetically favoured over the reformation of the intermolecular stem. Duplex Scorpions have significant advantages over the stem-loop version counterpart, producing a more intense fluorescence signal, as the quencher oligonucleotide is released into solution thereby minimising residual quenching. This greatly simplifies the oligonucleotide probe design. Also, Duplex Scorpions are simpler to synthesise and to purify by HPLC than the stem-loop version, as the fluorophore(s) and the quencher(s) are in separate oligonucleotides.[55]

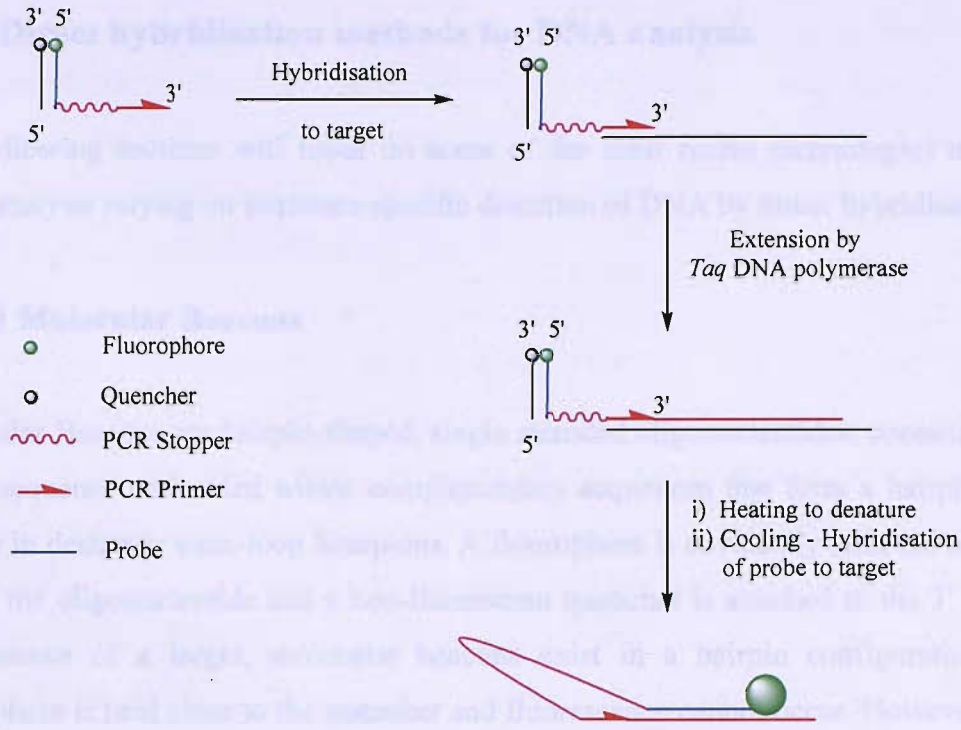


Figure 1.17 Schematics of a Duplex Scorpion detection system.

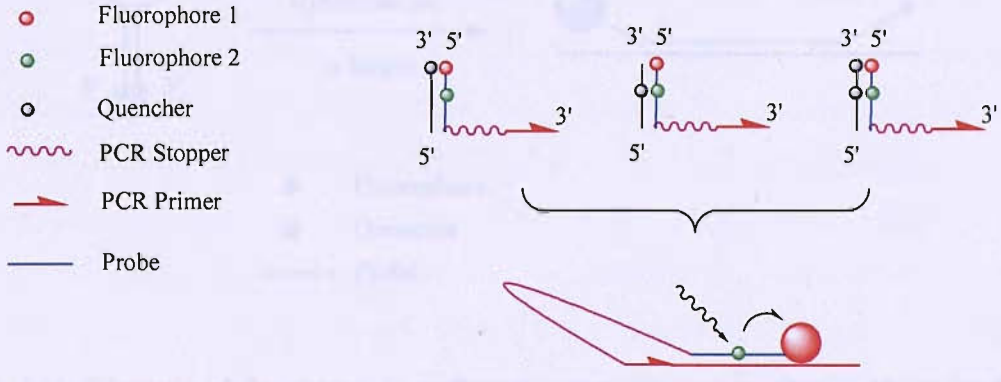


Figure 1.18 Schematics of a FRET duplex Scorpion detection system showing three different versions of the construct.

1.4.2 Direct hybridisation methods for DNA analysis

The following sections will focus on some of the most recent technologies used for DNA analysis relying on sequence specific detection of DNA by direct hybridisation.

1.4.2.1 Molecular Beacons

Molecular Beacons are hairpin-shaped, single stranded oligonucleotides, consisting of a probe sequence embedded within complementary sequences that form a hairpin stem, similar in design to stem-loop Scorpions. A fluorophore is covalently attached to the 5' end of the oligonucleotide and a non-fluorescent quencher is attached to the 3' end. In the absence of a target, molecular beacons exist in a hairpin configuration. The fluorophore is held close to the quencher and fluorescence cannot occur. However, once the beacon hybridises with target, the fluorophore and quencher are separated because the hairpin structure is disrupted, and fluorescence is detectable (Figure 1.19).[56, 57]

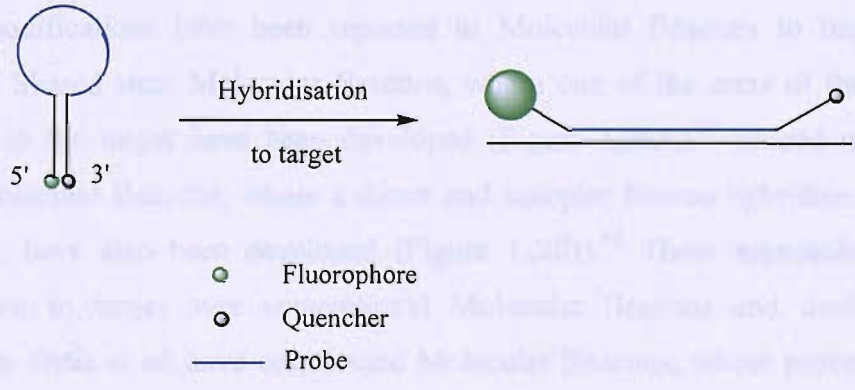


Figure 1.19 Schematic of the change in conformation that takes place when a Molecular Beacon hybridises with its specific target sequence. The blue loop portion of the beacon is a probe complementary to the target sequence in the nucleic acid to be detected.

Molecular Beacons permit a multiplex approach as they can detect a number of different targets in the same solution.[58] This is accomplished by constructing a different Molecular Beacon for each target and attaching a different fluorophore to each Molecular Beacon. The probes are mixed with the solution containing targets and the

colour that develops indicates which targets are present. Compared to linear probes, Molecular Beacons can have better specificity in gene detection; they can better discriminate single base-pair mismatches.[57, 59] For genotyping alleles, two Molecular Beacons differently labelled are used; one specific for the wild type allele and another specific for the mutant allele.[56]

When the probe encounters the target molecule it forms a probe target hybrid that is longer and more stable than the stem hybrid. The rigidity and length of the probe-target hybrid prevents the simultaneous existence of the stem duplex. Consequently, the Molecular Beacon undergoes a spontaneous conformational reorganisation that forces the stem hybrid to unfold and the fluorophore and quencher to move away from each other. However, designing Molecular Beacons for a specific application must be carefully studied as variations in probe loop or stem lengths can have a serious effect on hybridisation specificity and kinetics.[59] Depending on the length of the probe or the stems, the reformation of the intramolecular hairpin structure may be kinetically and entropically favoured over the intermolecular probe/target duplex formation.[59]

Various modifications have been reported to Molecular Beacons to improve their sensitivity. Shared stem Molecular Beacons, where one of the arms of the stem also hybridises to the target have been developed (Figure 1.20A).[60] Shared stem energy transfer Molecular Beacons, where a donor and acceptor beacon hybridise adjacent to each other, have also been developed (Figure 1.20B).[61] These approaches improve hybridisation to target over conventional Molecular Beacons and disfavour stem reformation. Ortiz *et al.* have constructed Molecular Beacons, whose probes contain a DNA analogue, PNA (Peptide nucleic acid).[62] PNA in general has a greater affinity for DNA than DNA or RNA.[63] These PNA Molecular Beacons have been designed for surface immobilisation to exploit Molecular Beacons using array formats.

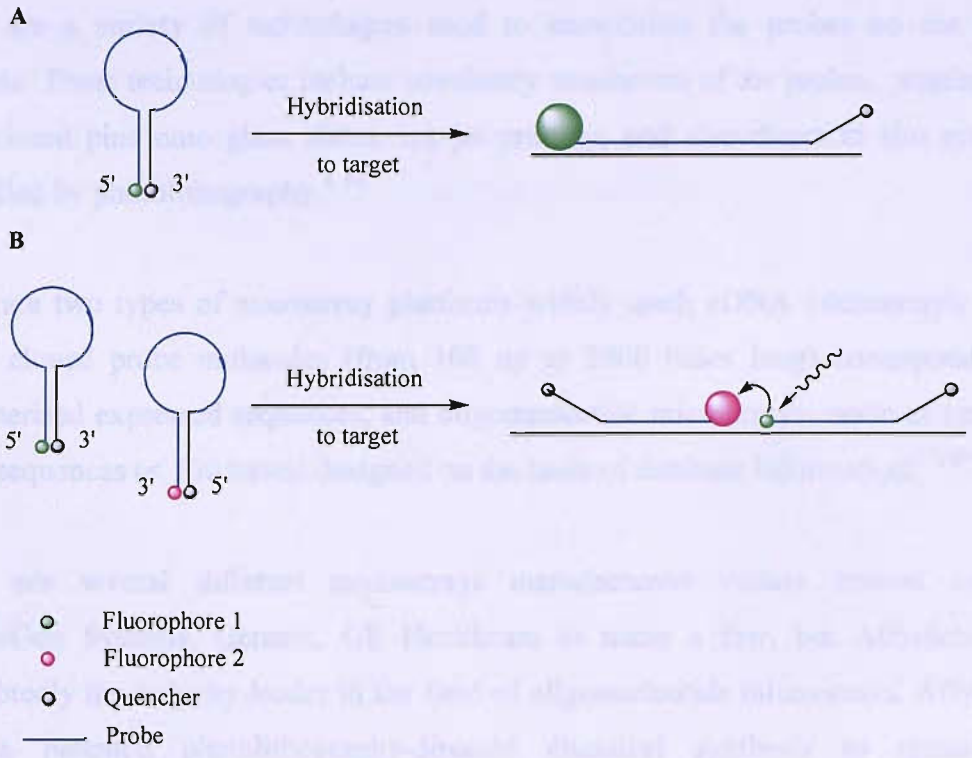


Figure 1.20 Shared stem Molecular Beacons (A) and shared stem energy transfer Molecular Beacons (B).

1.4.2.2 DNA planar microarrays

DNA microarrays have revolutionised biological and medical research by providing high throughput tools for DNA sequencing and the technology to monitor the expression of thousands of genes simultaneously and the capacity to analyse SNPs.[49, 50, 64-66] A DNA microarray-based analysis is a complex multistep process involving numerous specific pieces of equipments, and requiring expertise in a range of areas, including molecular biology, image analysis, computing, and statistics. There are several names for this technology - DNA microarrays, DNA arrays, DNA chips, gene chips, etc.[67]

Generically, planar microarrays are miniaturised devices consisting of hundreds to thousands of different oligonucleotide molecules, named "probes", that are orderly arranged on a microscopic scale onto a solid support such as a membrane or a glass microscope slide. The multiplexing capabilities of microarrays are produced by spatially encoding the array; each location on the array is used as a reporter of a specific target.[68]

There are a variety of technologies used to immobilise the probes on the planar supports. These technologies include covalently attachment of the probes, printing with fine-pointed pins onto glass slides, ink-jet printing, and also direct in situ synthesis controlled by photolithography.[4, 69]

There are two types of microarray platforms widely used; cDNA microarrays, which utilise cloned probe molecules (from 100 up to 2000 bases long) corresponding to characterised expressed sequences, and oligonucleotide microarrays, made of synthetic probe sequences (< 100 bases) designed on the basis of database information.[67, 69]

There are several different microarrays manufacturers widely known such as NimbleGen Systems, Genetix, GE Healthcare to name a few, but Affymetrix[70] is undoubtedly the industry leader in the field of oligonucleotide microarrays. Affymetrix uses a patented photolithography-directed chemical synthesis to manufacture GeneChip® Arrays, microarrays composed of hundreds of thousands of different oligonucleotides on a derivatised glass surface[69, 71] (Figure 1.21).

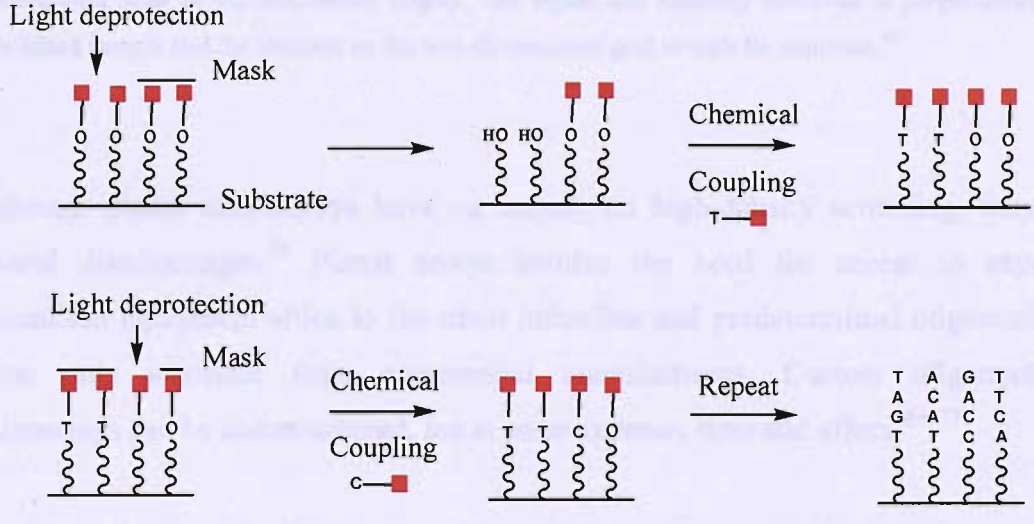


Figure 1.21 Probe synthesis on an Affymetrix GeneChip® Array, adapted from ref. 71. Lithographic masks are used to either block or transmit light onto specific locations of the substrate (chip) surface, leading to removal of protecting groups. The surface is then washed with a solution containing one of the nucleotides and coupling occurs only in those regions on the chip that have been deprotected through illumination. In this way, the microarray is built as the probes are synthesised through repeated cycles of deprotection and coupling. The process is repeated until the probes reach their full length.

Microarrays rely on the hybridisation between the immobilised probes and the complementary fluorescently labelled targets present in complex molecular mixtures[68] (Figure 1.22). The complementary hybridisation generates signals that reveal the identity and the concentration of the interacting labelled targets.[67, 70]

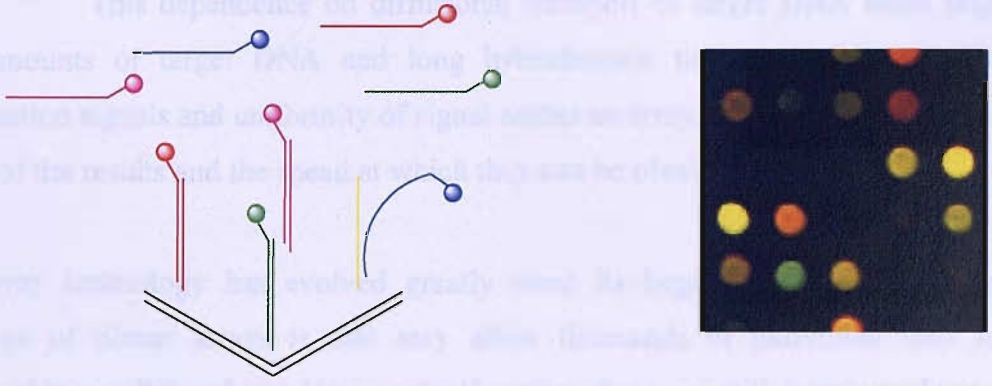


Figure 1.22 Microarray schematics showing the complementary hybridisation between the immobilised probes and the complementary labelled targets present in complex molecular pool (left). Non-complementary targets non-specifically bound to the probes (blue) are removed in the washes. Fluorescence scan of the microarray (right). The signal and intensity observed is proportional to the hybridised sample and the location on the two-dimensional grid reveals its sequence.[68]

Although planar microarrays have an impact on high-density screening, they have several disadvantages.[70] Planar arrays involve the need for access to expensive specialised equipment allied to the often inflexible and predetermined oligonucleotide chips only available from commercial manufacturers. Custom oligonucleotide microarrays can be commissioned, but at great expense, time and effort. [69, 72]

There are also several disadvantages depending on the array fabrication process.[68, 73] Probes are either spotted or printed to each position of the planar array individually or they are laboriously synthesised one nucleotide at a time. Both alternatives are time consuming. Consequently, there is a relatively low upper number to the number of arrays that can be produced at the same time. Another problem arises when the probes are covalently immobilised. They are all attached to the array under the same conditions using the same surface chemistry, which might not be suitable for all of the probes.

Also, the surface probe density as well as its spacing from the surface might have an impact on the extent to which the immobilised probes are able to hybridise to their targets from solution. The immobilised probes can be sterically inaccessible.[68]

Another major disadvantage stems from the feature properties of the planar surface; the rates of hybridisation are limited by slow diffusional kinetics of the targets to the surface.[68, 73] This dependence on diffusional transport of target DNA often requires large amounts of target DNA and long hybridisation times to achieve sufficient hybridisation signals and uniformity of signal across an array.[74] This severely affects the quality of the results and the speed at which they can be obtained.

Microarray technology has evolved greatly since its beginning. Perhaps, the main advantage of planar arrays is that they allow thousands of individual tests to be performed in parallel and in a large-scale. However, there are still many problems with the design of the platform which need to be addressed and overcome.[67, 69]

1.4.2.3 DNA suspension arrays: evolution of the planar microarray

DNA suspension arrays are DNA detection platforms in which the oligonucleotide probes are immobilised[75, 76] or synthesised[50] on micron-sized spherical beads.[49, 50, 75, 76] Microbeads have significant advantages over planar arrays in the terms of the way they are produced and used.[73] They are inexpensive, efficient and require mild operating conditions.[68] Oligonucleotide probes can be conjugated to millions of microbeads at the same time, with a degree of reproducibility that is impossible to achieve in planar array production. Also, different probes can be attached to different batches of microbeads, by a variety of chemistries and optimal conditions for each type of probes. This allows modification of the suspension platform by adding or subtracting beads with different probes. Planar arrays, however, need to be reprinted, requiring specialised equipment and more time.

Microbeads have the advantage of enhanced surface area compared to planar supports. The improvement in surface to volume ratio of the beads increases the probe density,

which in turn increases the binding capacity of the probes and the dynamic range of the assay. Also, the three-dimensional nature of the bead system, allows efficient mixing and close proximity between the beads and the targets, which accelerates the kinetics of target hybridisation.[68, 73, 76-78] This nearly solution-phase kinetics contrasts with planar arrays where conventional static hybridisations must allow sufficient time, typically 14–24 hours, for diffusion driven by random thermal motion to deliver targets to complementary probes.

Microbeads also facilitate the separation and washing steps and may even allow these to be eliminated altogether.[73] Another advantage is that microbeads allow minute sample volumes to be interrogated. This reduces the amount of valuable target sample and reagents required for the simultaneous detection and measurement of multiple reactions in a single vial. This ability to perform highly parallel throughput measurements is vital for maximising the efficiency and speed of the analysis. However, the multiplex format of the suspension array depends on the ability to decode the specific identity of the probe on any particular microbead. Thus, microbead assays have been used mainly in low density array applications.[68] To overcome this limitation, several encoding technologies have been developed in the last few years in order to expand the multiplex capabilities of suspension arrays.[68, 73, 77] These encoding technologies range from physical to optical; the former involves microfabrication of particles with physical different shapes or sizes,[73] the latter embeds the microparticles with fluorescent quantum dots[79-81] or dyes.[78, 82]

Luminex Corporation,[82] one of the leading companies in the field of suspension arrays, developed the *xMAP technology*, a system that uses fluorescence encoding technology to distinguish between solid supports. Their system uses micron sized polystyrene microspheres that are internally dyed with two spectrally distinct red and infra-red fluorochromes. Using precise amounts of each of these fluorochromes, an array is created consisting of up to 100 different beads sets with specific spectral addresses (Figure 1.23). Therefore an immobilised probe is identified by a specific spectral address. This allows up to 100 different targets to be screened simultaneously in a single reaction vessel.[78, 82] Armstrong *et al.*[75] have demonstrated the use of the *xMAP technology* on a multiplexed SNP genotyping assay (Figure 1.24).

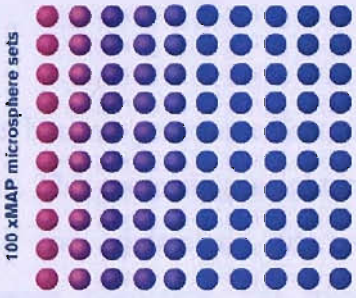


Figure 1.23 Luminex *xMAP technology*. Adapted from ref. 82.

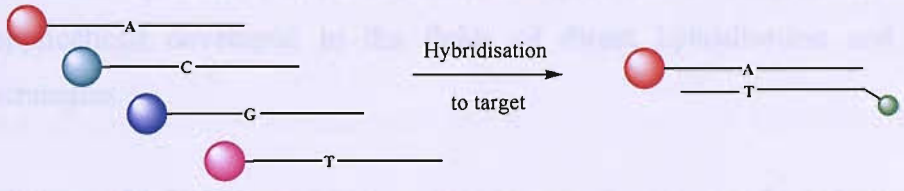


Figure 1.24 Multiplexed SNP genotyping assay using a suspension array. Flow cytometry analysis of the beads allows identification of the complementary hybridisation. Adapted from ref. 75.

The versatility of suspension arrays extends to immuno assays, proteins, lipids or carbohydrates.[77] The combination of highly parallel sample analysis and rapid serial delivery of samples makes suspension array technology an extremely attractive platform for a range of large-scale biomedical applications.

1.5 Conclusions

The analysis of SNPs in the human genome offers the key to understanding the significance of genetic variation. This promising insight into the history of human genes and the potential to reveal and understand the relationship between genetic information/variation and disease phenotype has triggered a massive technology revolution.

Novel technologies are constantly being explored to meet the urgent need for fast, inexpensive and reliable DNA analysis on a large scale. This chapter has presented a selection of technologies for DNA analysis focusing on some of the most recent and relevant applications developed in the fields of direct hybridisation and enzyme mediated strategies.

All the techniques for DNA analysis described possess key properties, advantages and disadvantages as referred throughout the chapter and perhaps because of that, no methodology for identifying SNPs has become accepted as a model technique and sometimes the task of choosing an analysis method may be difficult. So far, cost and throughput are the key limitations of sequencing and the challenge to develop the ideal methodology continues.[83]

1.6 References

1. Chargaff, E.; Vischer, E.; Doniger, R.; Green, C.; Misani, F., The composition of the desoxypentose nucleic acids of thymus and spleen. *J. Biol. Chem.* **1949**, 177, 405-16.
2. Crick, F., The Complementary Structure of DNA. *Proc. Natl. Acad. Sci. USA* **1954**, 40, 756-8.
3. Watson, J.D.; Crick, F.H., Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* **1953**, 171, 737-8.
4. Franklin, R.E.; Gosling, R.G., Molecular configuration in sodium thymonucleate. *Nature* **1953**, 171, 740-1.
5. Franklin, R.E.; Gosling, R.G., Evidence for 2-chain helix in crystalline structure of sodium deoxyribonucleate. *Nature* **1953**, 172, 156-7.
6. http://en.wikipedia.org/wiki/Image:ADN_animation.gif.
7. Mir, K.U.; Southern, E.M., Determining the influence of structure on hybridization using oligonucleotide arrays. *Nat. Biotechnol.* **1999**, 17, 788-92.

8. Southern, E.; Mir, K.U.; Shchepinov, M., Molecular interactions on microarrays. *Nat. Genet.* **1999**, 21, 5-9.
9. Strachan, T.; Read, A.P., Human Molecular Genetics 2. *Bios* **1999**, John Wiley and Sons Inc., New York, (2ed).
10. Dorris, D.R.; Nguyen, A.; Gieser, L.; Lockner, R.; Lublinsky, A.; Patterson, M.; Touma, E.; Sendera, T.J.; Elghanian, R.; Mazumder, A., Oligodeoxyribonucleotide probe accessibility on a three-dimensional DNA microarray surface and the effect of hybridization time on the accuracy of expression ratios. *BMC Biotechnol.* **2003**, 3, 6-17.
11. Owczarzy, R., Predictions of short DNA duplex thermodynamics and evaluation of next nearest neighbor interactions. *Ph.D. - University of Illinois* **1999**.
12. Watterson, J.; Piunno, P.A.E.; Krull, U.J., Practical physical aspects of interfacial nucleic acid oligomer hybridisation for biosensor design. *Anal. Chim. Acta* **2002**, 469, 115-27.
13. Eichhorn, G. L.; Shin, Y.A., Interaction of metal ions with polynucleotides and related compounds. XII. The relative effect of various metal ions on DNA helicity. *J. Am. Chem. Soc.* **1968**, 90, 7323-8.
14. SantaLucia, J.Jr.; Allawi, H.T.; Seneviratne, P.A., Improved nearest-neighbor parameters for predicting DNA duplex stability. *Biochemistry* **1996**, 35, 3555-62.
15. Allawi, H.T.; SantaLucia, J.Jr., Nearest-neighbor thermodynamics of internal A.C mismatches in DNA: sequence dependence and pH effects. *Biochemistry* **1998**, 37, 9435-44.
16. Carlson, C.S.; Newman, T.L.; Nickerson, D.A., SNPing in the human genome. *Curr. Opin. Chem. Biol.* **2001**, 5, 78-85.
17. Thelwell, N.; Millington, S.; Solinas, A.; Booth, J.; Brown, T., Mode of action and application of Scorpion primers to mutation detection. *Nucleic Acids Res.* **2000,** 28, 3752-61.
18. Haff, L.A.; Smirnov, I.P., Single-nucleotide polymorphism identification assays using a thermostable DNA polymerase and delayed extraction MALDI-TOF mass spectrometry. *Genome Res.* **1997**, 7, 378-88.
19. Nikiforov, T.T.; Rendle, R.B.; Goelet, P.; Rogers, Y.H.; Kotewicz, M.L.; Anderson, S.; Trainor, G.L.; Knapp, M.R., Genetic Bit Analysis: a solid phase method for typing single nucleotide polymorphisms. *Nucleic Acids Res.* **1994**, 22, 4167-75.
20. Ranasinghe, R.T., Novel techniques for genetic analysis. *Ph.D. - University of Southampton* **2004**.
21. Stoneking, M., Single nucleotide polymorphisms. From the evolutionary past. *Nature* **2001**, 409, 821-2.
22. Chakravarti, A., To a future of genetic medicine. *Nature* **2001**, 409, 822-3.
23. Sachidanandam, R.; Weissman, D.; Schmidt, S.C.; Kakol, J.M.; Stein, L.D.; Marth, G.; Sherry, S.; Mullikin, J.C.; Mortimore, B.J.; Willey, D.L.; Hunt, S.E.; Cole, C.G.; Coggill, P.C.; Rice, C.M.; Ning, Z.; Rogers, J.; Bentley, D.R.; Kwok, P.Y.; Mardis, E.R.; Yeh, R.T.; Schultz, B.; Cook, L.; Davenport, R.; Dante, M.; Fulton, L.; Hillier, L.; Waterston, R.H.; McPherson, J.D.; Gilman, B.; Schaffner, S.; Van Etten, W.J.; Reich, D.; Higgins, J.; Daly, M.J.; Blumenstiel, B.; Baldwin, J.; Stange-Thomann, N.; Zody, M.C.; Linton, L.; Lander, E.S.; Altshuler, D., A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature* **2001**, 409, 928-33.
24. Rockenbauer, E.; Peterson, K.; Vogel, U.; Bolund, L.; Kolvraa, S.; Nielsen, K.V.; Nexo, B.A., SNP genotyping using microsphere-linked PNA and flow cytometric detection. *Cytometry A* **2005**, 64, 80-6.

25. Mullis, K.B., Target amplification for DNA analysis by the polymerase chain reaction. *Ann. Biol. Clin. (Paris)* **1990**, 48, 579-82.
26. Mullis, K.B., The unusual origin of the polymerase chain reaction. *Sci. Am.* **1990**, 262, 56-61, 64-5.
27. Pati, N.; Schowinsky, V.; Kokanovic, O.; Magnuson, V.; Ghosh, S., A comparison between SNaPshot, pyrosequencing, and biplex invader SNP genotyping methods: accuracy, cost, and throughput. *J. Biochem. Biophys. Methods* **2004**, 60, 1-12.
28. Shapero, M.H.; Leuther, K.K.; Nguyen, A.; Scott, M.; Jones, K.W., SNP genotyping by multiplexed solid-phase amplification and fluorescent minisequencing. *Genome Res.* **2001**, 11, 1926-34.
29. Rosenblum, B.B.; Lee, L.G.; Spurgeon, S.L.; Khan, S.H.; Menchen, S.M.; Heiner, C.R.; Chen, S.M., New dye-labeled terminators for improved DNA sequencing patterns. *Nucleic Acids Res.* **1997**, 25, 4500-4.
30. Sanger, F.; Nicklen, S.; Coulson, A.R., DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **1977**, 74, 5463-7.
31. Sanger, F., Determination of nucleotide sequences in DNA. *Science* **1981**, 214, 1205-10.
32. Smith, L.M.; Fung, S.; Hunkapiller, M.W.; Hunkapiller, T.J.; Hood, L.E., The synthesis of oligonucleotides containing an aliphatic amino group at the 5' terminus: synthesis of fluorescent DNA primers for use in DNA sequence analysis. *Nucleic Acids Res.* **1985**, 13, 2399-412.
33. Voss, H.; Schwager, C.; Wirkner, U.; Sproat, B.; Zimmermann, J.; Rosenthal, A.; Erfle, H.; Stegemann, J.; Ansorge, W., Direct genomic fluorescent on-line sequencing and analysis using in vitro amplification of DNA. *Nucleic Acids Res.* **1989**, 17, 2517-27.
34. Prober, J.M.; Trainor, G.L.; Dam, R.J.; Hobbs, F.W.; Robertson, C.W.; Zagursky, R.J.; Cocuzza, A.J.; Jensen, M.A.; Baumeister, K., A system for rapid DNA sequencing with fluorescent chain-terminating dideoxynucleotides. *Science* **1987**, 238, 336-41.
35. http://en.wikipedia.org/wiki/Image:Sanger_sequencing_read_display.gif.
36. Cai, H.; White, P.S.; Torney, D.; Deshpande, A.; Wang, Z.L.; Keller, R.A.; Marrone, B.; Nolan, J.P., Flow cytometry-based minisequencing: A new platform for high- throughput single-nucleotide polymorphism scoring. *Genomics* **2000**, 66, 135-43.
37. Syvanen, A.C.; Aalto-Setala, K.; Harju, L.; Kontula, K.; Soderlund, H., A primer-guided nucleotide incorporation assay in the genotyping of apolipoprotein E. *Genomics* **1990**, 8, 684-92.
38. Curcio, M.; Stalhandske, P.; Lindberg, P.; Roeraade, J., Multiplex high-throughput solid-phase minisequencing by capillary electrophoresis and liquid core waveguide fluorescence detection. *Electrophoresis* **2002**, 23, 1467-72.
39. Landegren, U.; Kaiser, R.; Sanders, J.; Hood, L., A ligase-mediated gene detection technique. *Science* **1988**, 241, 1077-80.
40. Barany, F., Genetic disease detection and DNA amplification using cloned thermostable ligase. *Proc. Natl. Acad. Sci. USA* **1991**, 88, 189-93.
41. Landegren, U.; Schallmeiner, E.; Nilsson, M.; Fredriksson, S.; Baner, J.; Gullberg, M.; Jarvius, J.; Gustafsdottir, S.; Dahl, F.; Soderberg, O.; Ericsson, O.; Stenberg, J., Molecular tools for a molecular medicine: analyzing genes, transcripts and proteins using padlock and proximity probes. *J. Mol. Recognit.* **2004**, 17, 194-7.
42. Chen, X.; Livak, K.J.; Kwok, P.Y., A homogeneous, ligase-mediated DNA diagnostic test. *Genome Res.* **1998**, 8, 549-56.

43. Chen, X.; Kwok, P.Y., Homogeneous genotyping assays for single nucleotide polymorphisms with fluorescence resonance energy transfer detection. *Genet. Anal.* **1999**, 14, 157-63.
44. LaRussa, P.; Lungu, O.; Hardy, I.; Gershon, A.; Steinberg, S.P.; Silverstein, S., Restriction fragment length polymorphism of polymerase chain reaction products from vaccine and wild-type varicella-zoster virus isolates. *J. Virol.* **1992**, 66, 1016-20.
45. Ronai, Z.; Sasvari-Szekely, M.; Guttman, A., Miniaturized SNP detection: quasi-solid-phase RFLP analysis. *Biotechniques* **2003**, 34, 1172-3.
46. Liu, W.H.; Kaur, M.; Makrigiorgos, G.M., Detection of hotspot mutations and polymorphisms using an enhanced PCR-RFLP approach. *Hum. Mutat.* **2003**, 21, 535-41.
47. Huber, S.; McMaster, K.J.; Voelkerding, K.V., Analytical evaluation of primer engineered multiplex polymerase chain reaction-restriction fragment length polymorphism for detection of factor V Leiden and prothrombin G20210A. *J. Mol. Diagn.* **2000**, 2, 153-7.
48. Kaur, M.; Zhang, Y.; Liu, W.H.; Tetradis, S.; Price, B.D.; Makrigiorgos, G.M., Ligation of a primer at a mutation: a method to detect low level mutations in DNA. *Mutagenesis* **2002**, 17, 365-74.
49. Brenner, S.; Johnson, M.; Bridgham, J.; Golda, G.; Lloyd, D.H.; Johnson, D.; Luo, S.; McCurdy, S.; Foy, M.; Ewan, M.; Roth, R.; George, D.; Eletr, S.; Albrecht, G.; Vermaas, E.; Williams, S.R.; Moon, K.; Burcham, T.; Pallas, M.; DuBridge, R.B.; Kirchner, J.; Fearon, K.; Mao, J.; Corcoran, K., Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. *Nat. Biotechnol.* **2000**, 18, 630-4.
50. Brenner, S.; Williams, S.; Vermaas, E.H.; Storck, T.; Moon, K.; McCollum, C.; Mao, J.I.; Luo, S.; Kirchner, J.J.; Eletr, S.; DuBridge, R.B.; Burcham, T.; Albrecht, G., In vitro cloning of complex mixtures of DNA on microbeads: physical separation of differentially expressed cDNAs. *Proc. Natl. Acad. Sci. USA* **2000**, 97, 1665-70.
51. Che, Y.; Chen, X., A multiplexing single nucleotide polymorphism typing method based on restriction-enzyme-mediated single-base extension and capillary electrophoresis. *Anal. Biochem.* **2004**, 329, 220-9.
52. Bruland, O.; Knappskog, P.M., One-tube restriction enzyme digest and fluorescent labeling for restriction endonuclease fingerprinting single-strand conformational polymorphism. *Biotechniques* **2004**, 37, 906-12.
53. Whitcombe, D.; Theaker, J.; Guy, S.P.; Brown, T.; Little, S., Detection of PCR products using self-probing amplicons and fluorescence. *Nat. Biotechnol.* **1999**, 17, 804-7.
54. McKeen, C.M.; Brown, L.J.; Nicol, J.T.; Mellor, J.M.; Brown, T., Synthesis of fluorophore and quencher monomers for use in scorpion primers and nucleic acid structural probes. *Org. Biomol. Chem.* **2003**, 1, 2267-75.
55. Solinas, A.; Brown, L.J.; McKeen, C.; Mellor, J.M.; Nicol, J.; Thelwell, N.; Brown, T., Duplex Scorpion primers in SNP analysis and FRET applications. *Nucleic Acids Res.* **2001**, 29, e96.
56. Kostrikis, L.G.; Tyagi, S.; Mhlanga, M.M.; Ho, D.D.; Kramer, F.R., Spectral genotyping of human alleles. *Science* **1998**, 279, 1228-9.
57. Tyagi, S.; Bratu, D.P.; Kramer, F.R., Multicolor molecular beacons for allele discrimination. *Nat. Biotechnol.* **1998**, 16, 49-53.
58. Marras, S.A.; Kramer, F.R.; Tyagi, S., Multiplex detection of single-nucleotide variations using molecular beacons. *Genet. Anal.* **1999**, 14, 151-6.

59. Tsourkas, A.; Behlke, M.A.; Rose, S.D.; Bao, G., Hybridization kinetics and thermodynamics of molecular beacons. *Nucleic Acids Res.* **2003**, 31, 1319-30.
60. Tsourkas, A.; Behlke, M.A.; Bao, G., Structure-function relationships of shared-stem and conventional molecular beacons. *Nucleic Acids Res.* **2002**, 30, 4208-15.
61. Tsourkas, A.; Behlke, M.A.; Xu, Y.; Bao, G., Spectroscopic features of dual fluorescence/luminescence resonance energy-transfer molecular beacons. *Anal. Chem.* **2003**, 75, 3697-703.
62. Ortiz, E.; Estrada, G.; Lizardi, P.M., PNA molecular beacons for rapid detection of PCR amplicons. *Mol. Cell. Probes* **1998**, 12, 219-26.
63. Ratilainen, T.; Holmen, A.; Tuite, E.; Nielsen, P.E.; Norden, B., Thermodynamics of sequence-specific binding of PNA to DNA. *Biochemistry* **2000**, 26, 7781-91.
64. Dai, H.; Meyer, M.; Stepaniants, S.; Ziman, M.; Stoughton, R., Use of hybridization kinetics for differentiating specific from non-specific binding to oligonucleotide microarrays. *Nucleic Acids Res.* **2002**, 30, e86.
65. Lemeshko, S.V.; Powdrill, T.; Belosludtsev, Y.Y.; Hogan, M., Oligonucleotides form a duplex with non-helical properties on a positively charged surface. *Nucleic Acids Res.* **2001**, 29, 3051–8.
66. Riccelli, P.V.; Merante, F.; Leung, K.T.; Bortolin, S.; Zastawny, R.L.; Janeczko, R.; Benight, A.S., Hybridization of single-stranded DNA targets to immobilized complementary DNA probes: comparison of hairpin versus linear capture probes. *Nucleic Acids Res.* **2001**, 4, 996-1004.
67. Gershon, D., Microarray technology: an array of opportunities. *Nature* **2002**, 25, 885-91.
68. Situma, C.; Hashimoto, M.; Soper, S.A., Merging microfluidics with microarray-based bioassays. *Biomol. Eng.* **2006**, 23, 213-31.
69. Murphy, D., Gene expression studies using microarrays: principles, problems, and prospects. *Adv. Physiol. Educ.* **2002,** 26, 256-70.
70. http://www.affymetrix.com.
71. Lipshutz, R.J.; Fodor, S.P.; Gingeras, T.R.; Lockhart, D.J., High density synthetic oligonucleotide arrays. *Nat. Genet.* **1999**, 21, 20-4.
72. Yang, L.; Tran, D.K.; Wang X., BADGE, Beads Array for the Detection of Gene Expression, a high-throughput diagnostic bioassay. *Genome Res.* **2001**, 11, 1888-98.
73. Wilson, R.; Cossins, A.R.; Spiller, D.G., Encoded microcarriers for high-throughput multiplexed detection. *Angew. Chem. Int. Ed. Engl.* **2006**, 45, 6104-17.
74. McQuain, M.K.; Seale, K.; Peek, J.; Fisher, T.S.; Levy, S.; Stremler, M.A.; Haselton, F.R., Chaotic mixer improves microarray hybridization. *Anal. Biochem.* **2004**, 15, 215-26.
75. Armstrong, B.; Stewart, M.; Mazumder, A., Suspension arrays for high throughput, multiplexed single nucleotide polymorphism genotyping. *Cytometry* **2000**, 2, 102-8.
76. Spiro, A.; Lowe, M.; Brown, D., A bead-based method for multiplexed identification and quantitation of DNA sequences using flow cytometry. *Appl. Environ. Microbiol.* **2000**, 66, 4258-65.
77. Nolan, J.P.; Sklar, L.A., Suspension array technology: evolution of the flat-array paradigm. *Trends Biotechnol.* **2002**, 20, 9-12.
78. Dunbar, SA., Applications of Luminex xMAP technology for rapid, high-throughput multiplexed nucleic acid detection. *Clin. Chim. Acta* **2006**, 363, 71-82.

79. O'Brien, P.; Cummins, S.S.; Darcy, D.; Dearden, A.; Masala, O.; Pickett, N.L.; Ryley, S.; Sutherland, A.J., Quantum dot-labelled polymer beads by suspension polymerisation. *Chem. Commun. (Camb.)* **2003**, 20, 2532-3.
80. Gao, X.; Nie, S., Quantum dot-encoded beads. *Methods Mol. Biol.* **2005**, 303, 61-71.
81. Han, M.; Gao, X.; Su, J.Z.; Nie, S., Quantum-dot-tagged microbeads for multiplexed optical coding of biomolecules. *Nat. Biotechnol.* **2001**, 19, 631-5.
82. http://www.luminexcorp.com.
83. Bentley, D.R., Whole-genome re-sequencing. *Curr. Opin. Genet. Dev.* **2006**, 16, 545-52.

2. Initial experiments; Exploring probe attachment chemistry

2.1 4G project

The work described in this thesis is part of the basic technology grant (4G project) to design and implement a new method of synthesising, screening and sequencing DNA.[1] The 4G project seeks to analyse genetic variations on a global scale, at a rate up until now impossible. The rapid, accurate and comprehensive identification of all single base changes within the genotype of individuals will provide insight into the genetic variations with impact on disease phenotype.

The fundamental aim of the 4G project is the development of a practical encoding technology for micron-sized particles that can be used in combinatorial chemistry to generate very large encoded libraries of biomolecules. This technology can be applied to the production of very large libraries of DNA molecules that can be then used for sequence analysis; either for the detection of specific DNA sequences in a sample or ultimately for the detection of any sequence to enable ultra-high throughput DNA sequencing. The 4G project utilises the complementary hybridisation (self-recognition) of DNA and incorporation of a single fluorescently labelled ddNTP by a DNA polymerase as the basis of this sequencing technology (Figure 2.1).

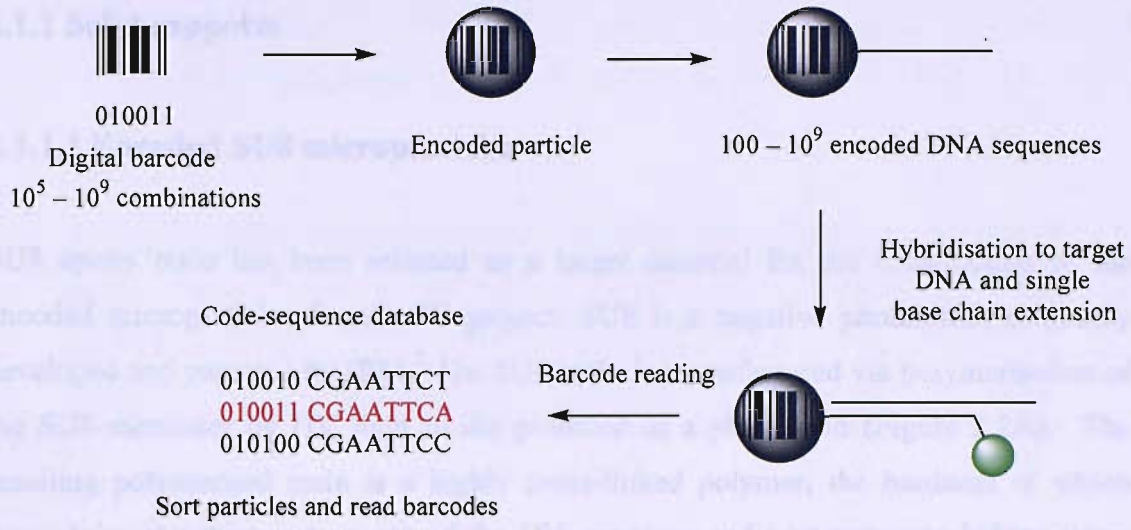


Figure 2.1 Schematic for the individual steps of the 4G sequencing project. Adapted from[1].

As Figure 2.2 shows, the 4G project aims to construct a probe library in which, each microparticle will have a single tag and a single oligonucleotide probe sequence attached. These probes might be specific to SNPs regions or even to PCR primers. This microparticle library will be mixed with and hybridised to a fluorescent target DNA sample. These targets may represent any sample ranging from single or multiple PCR products, to cDNA libraries, to complete amplified genomic DNA of one, or a number of individuals and to environmental DNA libraries. Hybridisation to target DNA will be followed by the incorporation of a single fluorescently labelled ddNTP by a DNA polymerase. Fluorescent microparticles will then be sorted according to fluorescence and the encoded tag will be read to identify the tethered DNA probe and the hybridised target DNA. It is important to note that only those beads with a fluorophore added by chain extension will need to be analysed. The sequencing and screening capabilities of the project depend on computational reassembly of the overlapping DNA contiguous sequences detected and decoded. The full scale of the probe and target sequence libraries produced, associated with the particle encoding technology will provide an ultra-high throughput suspension array.

2.1.1 Solid supports

2.1.1.1 Encoded SU8 microparticles

SU8 epoxy resin has been selected as a target material for the construction of the encoded microparticles for the 4G project. SU8 is a negative photoresist, originally developed and patented by IBM.[2] The SU8 resin is manufactured via polymerisation of the SU8 monomer by UV light in the presence of a photo acid (Figure 2.2A). The resulting polymerised resin is a highly cross-linked polymer, the hardness of which depends on the time and energy of the UV exposure and post-exposure baking times and temperatures.[3, 4] Photolithographic masks are used to pattern the SU8, producing well defined structures with excellent mechanical properties with high aspect ratio (ratio of height to width). Figure 2.2B shows a photolithografically fabricated un-encoded rectangular SU8 microparticle.

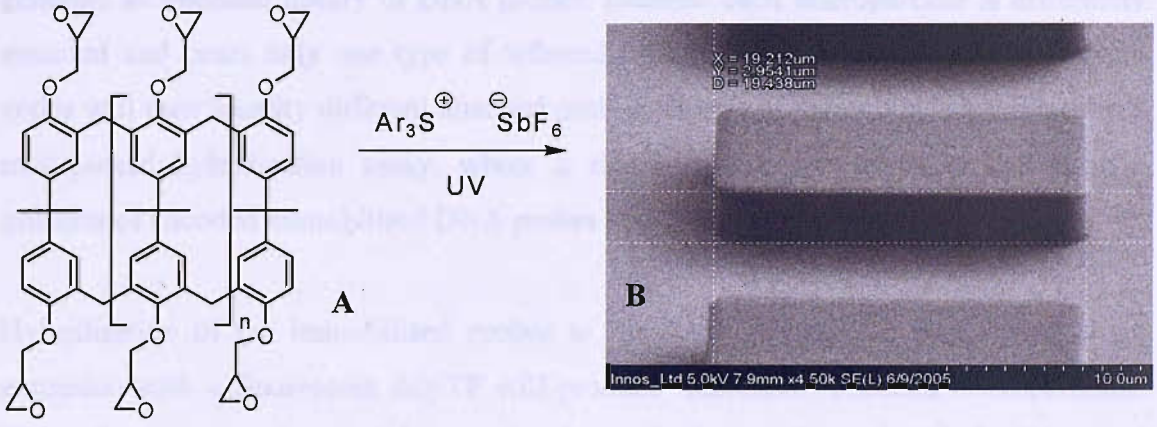


Figure 2.2 Photolithographic manufacturing of un-encoded SU8 microparticles (Scanning electron microscopy (SEM) on the right) after monomer polymerisation by UV light under acidic conditions.[3]

Different photolithographic masks can be used to pattern the SU8 rectangular microparticles differently, generating encoded particles as depicted in Figure 2.3.

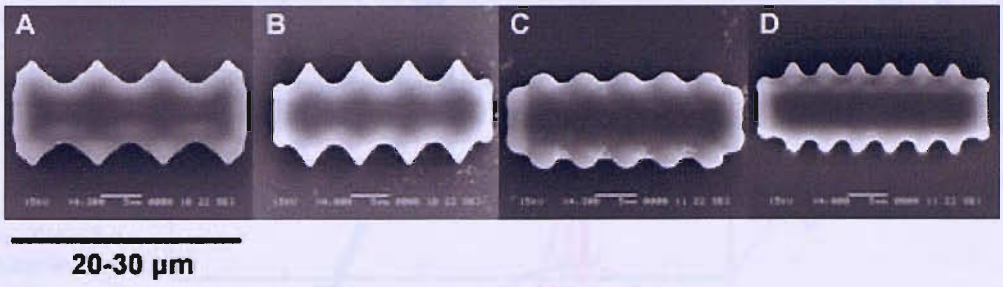


Figure 2.3 SEM of four differently encoded SU8 microparticles.

The free residual epoxy groups on the surface of the microparticles are then used for further chemistry, which may include direct synthesis or immobilisation of biomolecules. The attachment of oligonucleotides onto encoded SU8 microparticles will generate an encoded library of DNA probes. Because each microparticle is differently encoded and bears only one type of tethered probe, different particles with different codes will then identity different attached probes. This is of paramount importance in a multiplexed hybridisation assay, where a single solution environment can contain millions of encoded immobilised DNA probes and DNA targets.

Hybridisation of the immobilised probes to fluorescent targets or single base chain extension with a fluorescent ddNTP will produce fluorescent encoded microparticles. These fluorescent microparticles can be detected, their code read and finally sorted using an in-house developed detection system based on microfluidics[5] (Figure 2.4).

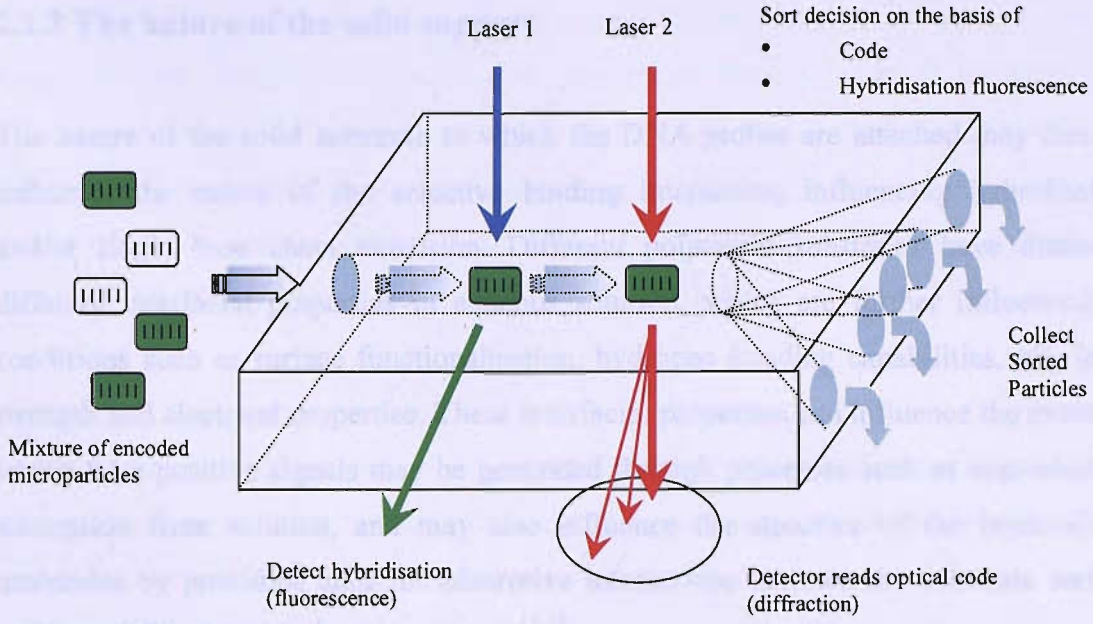


Figure 2.4 Schematic diagram of the in-house detection system. Laser 1 detects and measures the fluorescence of the particles. Laser 2 illuminates encoded microparticles producing a diffracting pattern and read the code. Coupling of both sources of detection enables unequivocal identification of the microparticles and the attached biomolecules.

2.1.1.2 Un-encoded commercially available solid supports

At the time of the experiments carried out in this chapter, the SU8 encoding technology was at an early phase. It was decided to explore commercially available solid supports to set up optimised conditions for the DNA probe attachment, hybridisation to target and single base chain extension. These processes would be transferred to SU8 microparticles at a later stage.

There are a wide range of commercially available solid supports described in the literature[6] for probe immobilisation, ranging from glass, ceramics, silicon wafer, magnetic beads, non-magnetic beads and nylon, to a variety of polymers and membranes. These solid supports vary from surfaces to discrete particles and to monodisperse beads, being the most commonly used: silica, glass and GMA (Glycidyl Methacrylate) [7-10] microparticles.

2.1.2 The nature of the solid support

The nature of the solid substrate to which the DNA probes are attached may directly influence the nature of the selective binding interaction, influencing hybridisation and/or single base chain extension. Different polymeric substrates have distinctly different interfacial properties in aqueous solution, which are further influenced by conditions such as surface functionalisation, hydrogen bonding capabilities, pH, ionic strength and electrical properties. These interfacial properties can influence the extent to which false positive signals may be generated through processes such as non-selective adsorption from solution, and may also influence the structure of the immobilised molecules by providing sites for adsorptive interactions between the substrate surface and immobilised molecules themselves. [11-15]

2.1.3 Chemistries for probe attachment and the immobilisation effect on hybridisation

There are a range of approaches[9, 10, 16-23] to immobilising DNA probes onto the surface of microparticles. These approaches differ in the nature of the chemical reactive groups on the surface of the supports and at the termini of the DNA probes. The nature of the chemical reaction driving the coupling varies also accordingly. The most widely used chemical combinations of reactive groups are the carbodiimide driven amide chemistry between NH_2 and COOH groups, [6, 9, 10, 20, 24] epoxy reaction with NH_2,[25, 26] and the avidin-biotin binding strategy. [21-23, 27] Also, chemical reaction between SH and NH_2,[16] OH and phosphate,[20] SH and gold,[28] and NH_2 and phosphate[29] have been explored.

The attachment of DNA to solid phase supports introduces problems which are not encountered when hybridisation takes place in a homogeneous solution as it constrains the way in which the probes and targets can interact with each other. The DNA probe is not so free to diffuse as it would be in solution, thus reducing rates of hybridisation. Target DNA is then prevented from making a close approach to the probe by a number of steric factors: it cannot make an approach from the direction of the solid support unless there is a spacer[11, 16, 30] between it and the immobilised probe; it may also be

hindered in its approach from the solution phase if the probes on the support are too close together. Close approach will also be affected by the nature of the surface to which the probes are attached; its charge, hydrophobicity and degree of solvation are all likely to influence significantly the environment in which the interaction occurs. The introduction of a spacer between solid supports and immobilised probes improves the coupling yield and forms a product of higher purity. [30]

While DNA hybridisation in solution is well studied, fewer investigations have focussed on the in-situ kinetics and thermodynamics for surface immobilised probes interacting with solution-phase targets, where the molecular processes are more complex. These processes relate to factors that are less well understood such as probe density, surface heterogeneity and non-specific adsorption. The importance of these complex factors is increasingly recognised but largely unaddressed in published experiments. [31] For example, probe density dependant kinetic or steric constraints may alter the apparent stability or selectivity of probe-target binding for surface immobilised oligonucleotides.

2.2 Aims of this chapter

This chapter will explore the covalent attachment of DNA probes onto solid supports, subsequent hybridisation to target and addition of a fluorescent ddNTP to the immobilised duplex by a DNA polymerase. Different chemistries of probe attachment will be studied.

An appropriate selection of un-encoded commercially available solid supports will be investigated. Un-encoded SU8 microparticles will also be explored. It is important that the solid supports used will permit probe coupling and hybridisation as well as enzyme accessibility to the immobilised construct.

2.3 Results and discussion

2.3.1 Choosing a solid support

The most commonly used commercially available solid supports referred in the literature, for oligonucleotide probe immobilisation are silica and GMA beads (section 2.1.1.2). They were considered as a starting point to optimising our on-bead DNA chemistry. SEM and Laser Diffraction Particle Size (LDPS) analysis for silica beads (5 and 10 μm) showed the size variability of the microparticles (Figure 2.5A and B). A forward scatter (FSC) versus side scatter (SSC 488/10-A) light intensity dot-plot obtained by FACS analysis (Figure 2.5C) also revealed a polydisperse population of silica beads (5 μm). A monodisperse bead preparation would be expected to appear as a single group of points or dots in the middle of the dot-plot, but the large variation in forward scatter indicated a wide distribution in the size of the silica beads.

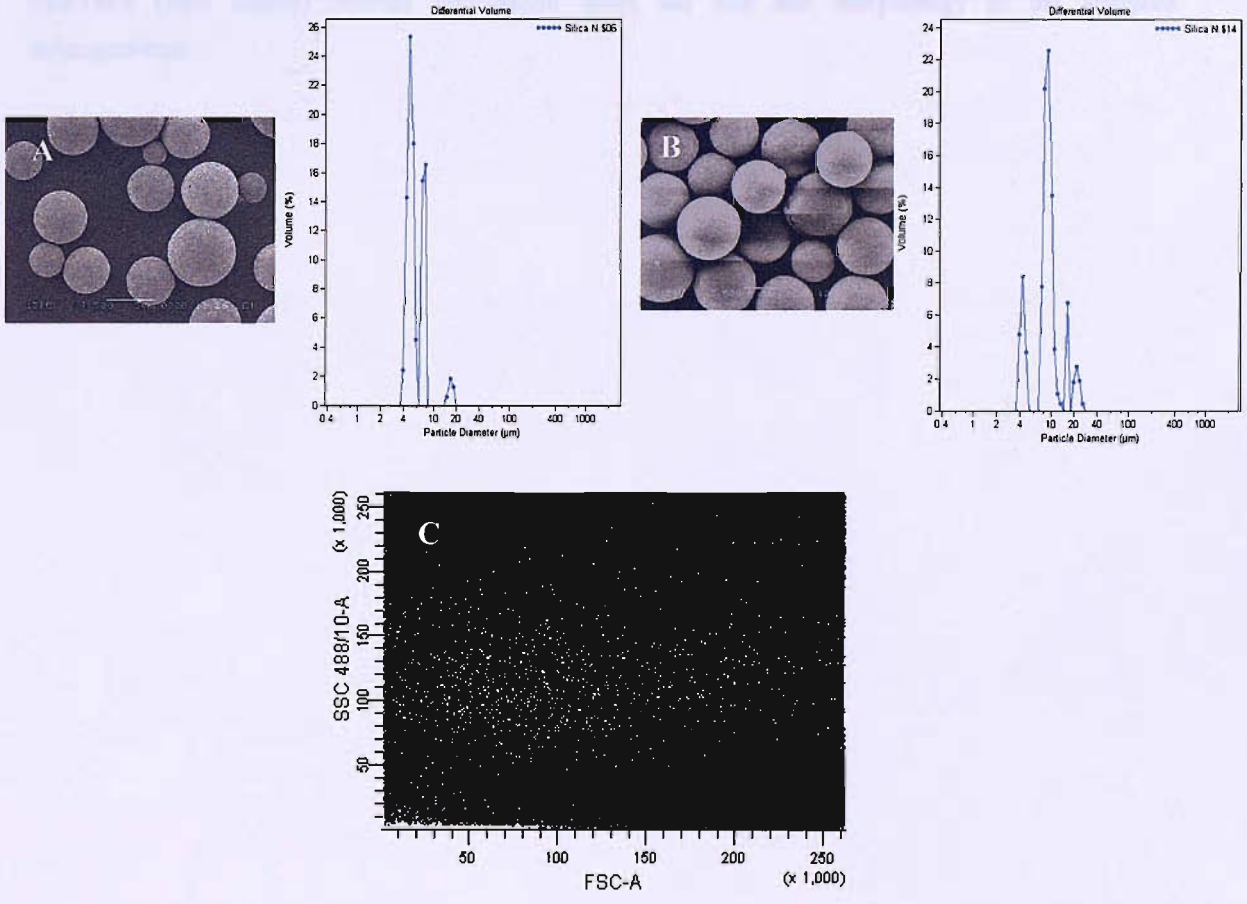


Figure 2.5 SEM and LDPS analysis of silica beads, 5 μm (A) and 10 μm (B). FACS analysis of silica beads, 5 μm (C); FSC-A (Forward scatter) and SSC 488/10-A (Side scatter) provide information about the size and morphology of the analysed microparticles.

On the other hand, GMA beads (5 μm), displayed a uniform population of microparticles, both by SEM and FACS analysis (Figures 2.6A and B). The red points on the centre of the dot-plot represent the main population (P1) of beads. The green points represent populations of aggregates of beads.

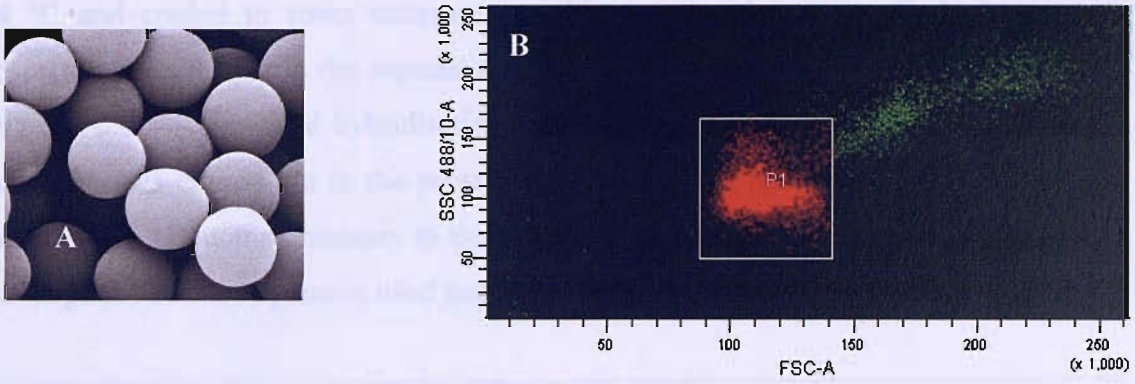


Figure 2.6 SEM (A) and FACS analysis (B) of GMA beads (5 μm); FSC-A (Forward scatter) and SSC 488/10-A (Side scatter) provide information about the size and morphology of the analysed microparticles.

2.3.2 Single base chain extension preliminary experiments

Single base chain extension preliminary studies, involved the attachment[32, 33] of a 5' NH_2 26-mer probe (4goligo0006) onto amino functionalised silica (5 μm) and GMA (5 μm) beads using glutaraldehyde (Figure 2.7A) and subsequent hybridisation to a complementary 30-mer target (4goligo0004). The hybridisation mixture was heated to 94 °C and cooled to room temperature with occasional agitation.[34] A decrease in absorbance at 260 nm in the supernatant of probe immobilisation and hybridisation to target reactions suggested hybridisation had occurred. Single base chain extension, or labelling, was carried out in the presence of *Taq* DNA polymerase and of fluorescein labelled ddATP, complementary to the next free base, T, in the target. Table 2.1 shows the oligonucleotide sequences used and Figure 2.7B outlines the experiment.

Oligonucleotide	Type of modification	Sequence5' to 3'	Length
Probe - 4goligo0006	5' NH_2	GAGATGCCCAGAATGAAGTCAAGTCG	26
Target - 4goligo0004	-	CCATCGACTTGACTTCATTCTGGGCATCTC	30

Table 2.1 Probe and target oligonucleotides used for preliminary labelling studies.

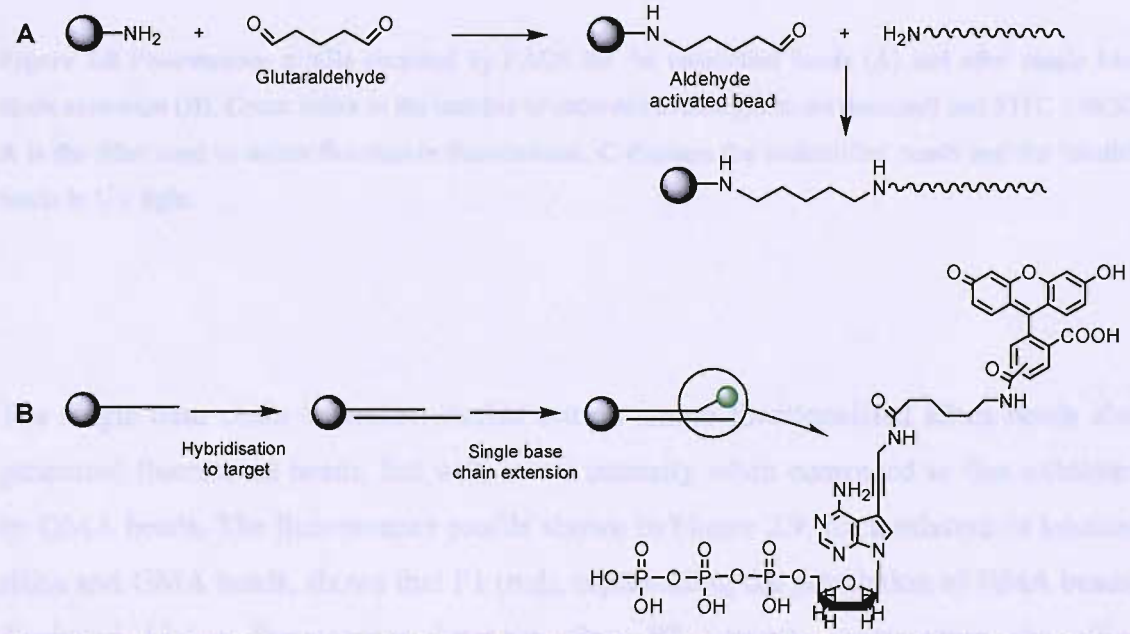


Figure 2.7 (A) Immobilisation of 5' NH_2 probes onto amino functionalised beads using glutaraldehyde. (B) Single base chain extension experiment outline and the structure of fluorescein – ddATP.

Figure 2.8 shows the fluorescence profile obtained by FACS for GMA beads after being washed. The beads exhibited a high increase in fluorescence after single base chain extension (Figure 2.8B), when compared to the signal displayed by the unmodified control (Figure 2.8A). Figure 2.8C shows the unmodified GMA control and the labelled GMA beads in UV light exposure.

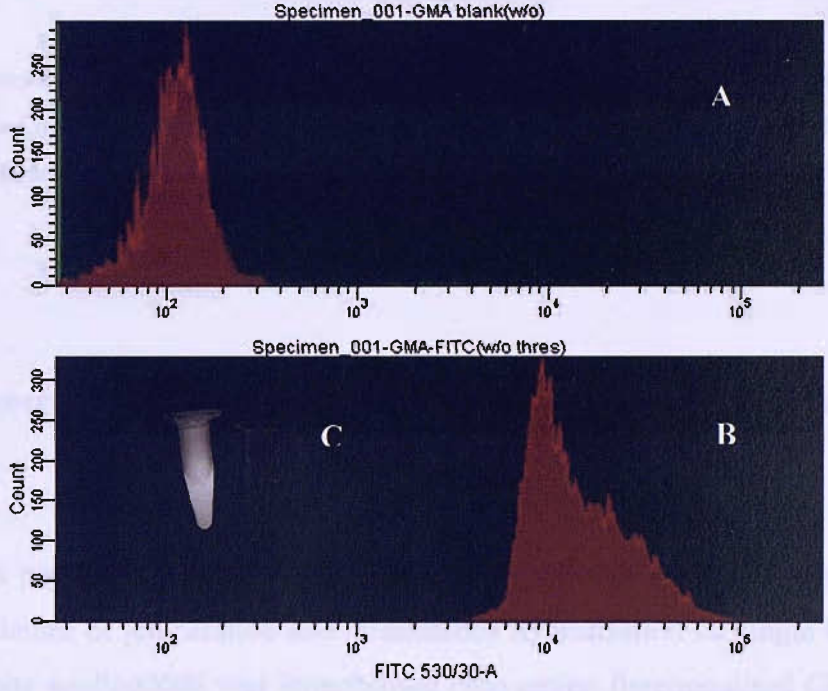


Figure 2.8 Fluorescence profile obtained by FACS for the unlabelled beads (A) and after single base chain extension (B). Count refers to the number of recorded events (particles detected) and FITC 530/30-A is the filter used to detect fluorescein fluorescence. C displays the unmodified beads and the labelled beads in UV light.

The single base chain extension carried out on amino functionalised silica beads also generated fluorescent beads, but with lower intensity when compared to that exhibited by GMA beads. The fluorescence profile shown in Figure 2.9, for a mixture of labelled silica and GMA beads, shows that P1 (red), representing the population of GMA beads, displayed higher fluorescence intensity than P2 (green), representing the silica population. Also, the silica population showed a higher spread in fluorescence, probably due to the wide distribution in particle size.

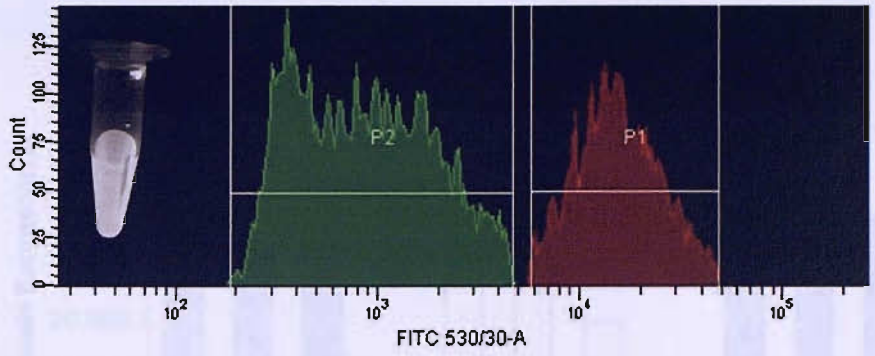


Figure 2.9 Fluorescence profile obtained by FACS for a mixture of labelled GMA (P1-red) and silica (P2-green) beads. Count refers to the number of recorded events (particles detected) and FITC 530/30-A is the filter used to detect fluorescein fluorescence. The microtube displays the labelled silica beads in UV light.

2.3.3 Influence of temperature and mismatches on single base chain extension

The seemingly positive results presented in previous section prompted an experiment to assess the influence of temperature and mismatched hybridisation on single base chain extension. Probe 4goligo0006 was immobilised onto amino functionalised GMA beads and hybridised to complementary (4golig0004) and mismatched (4goligo0011) targets (Table 2.2). Single base chain extension was carried out at 60, 70 and 80 °C on the complementary and mismatched duplexes. GMA beads with immobilised probe were incubated with *Taq* and fluorescein-ddATP to assess the level of non-specific binding.

Oligonucleotide	Type of modification	Sequence 5' to 3'	Length
Probe – 4goligo0006	5' NH_2	GAGATGCCCAGAATGAAGTCAAGTCG	26
Target – 4goligo0004	-	CCATCGACTTGACTTCATTCTGGGCATCTC	30
Target – 4goligo0011	-	CCATCGTCTTGACTTCATTCTGGGCATCTC	30

Table 2.2 Oligonucleotides used for the labelling. 4goligo0004 and 4goligo0011 are the complementary and mismatched targets respectively to immobilised probe 4goligo0006. The mismatched position is highlighted.

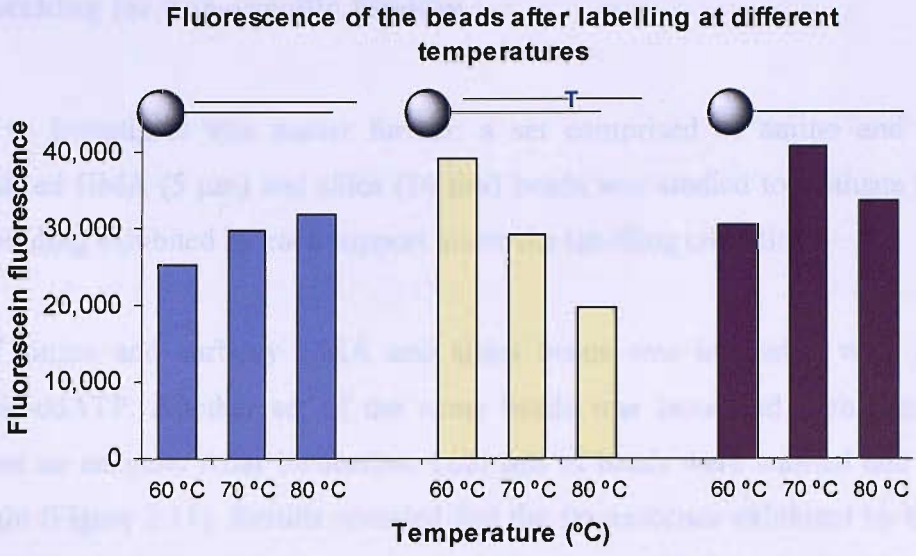


Figure 2.10 Fluorescence (± standard error (SE)) obtained by FACS for the labelling on the various controls at different temperatures. Fluorescence read-out of the beads is and will be presented throughout the thesis with the standard error associated.

It was anticipated that only the immobilised complementary duplexes would exhibit fluorescence. However, the results obtained by FACS displayed in Figure 2.10 show that all suspensions of beads exhibited high levels of fluorescence after single base chain extension at all the temperatures. The fact that the control of GMA beads with immobilised probe showed high levels of fluorescence was surprising. This suggested either a single strand chain extension by the enzyme or non-specific binding of the labelled nucleotide onto the support. *Taq* is known for the ability to add nucleotides in a template independent manner to the 3' ends of single strands.[35] To investigate the possibility of single strand extension, the experiment was repeated in the presence of a different DNA polymerase, *Pfu Turbo*, not prone to single strand extension and with higher nucleotide incorporation fidelity than *Taq*.[36] However, the fluorescence exhibited was the same for all suspensions. These findings suggested the high levels of fluorescence observed for the negative control were due to the non-specific binding of the labelled nucleotide onto the support rather than single strand chain extension. Moreover, the results suggested that the fluorescence observed in the positive controls could also be due to non-specific binding of the nucleotide.

2.3.4 Checking for non-specific binding

In order to investigate this matter further, a set comprised of amino and carboxy functionalised GMA (5 μm) and silica (10 μm) beads was studied to evaluate the non-specific binding exhibited by each support under the labelling conditions.

A set of amino and carboxy GMA and silica beads was incubated with *Taq* and fluorescein-ddATP. Another set of the same beads was incubated with fluorescein-ddATP but no enzyme. After incubation, both sets of beads were washed and exposed to UV light (Figure 2.11). Results revealed that the fluorescence exhibited by the set of beads incubated with *Taq* and fluorescein-ddATP (Figure 2.11A) is similar to that of the set of beads incubated with fluorescein-ddATP but no enzyme (Figure 2.11B). Figure 2.11C shows the intrinsic fluorescence of each type of unmodified beads for comparison. Results revealed that amino functionalised beads, both GMA and silica are the type most prone to non-specific binding of the labelled ddATP.

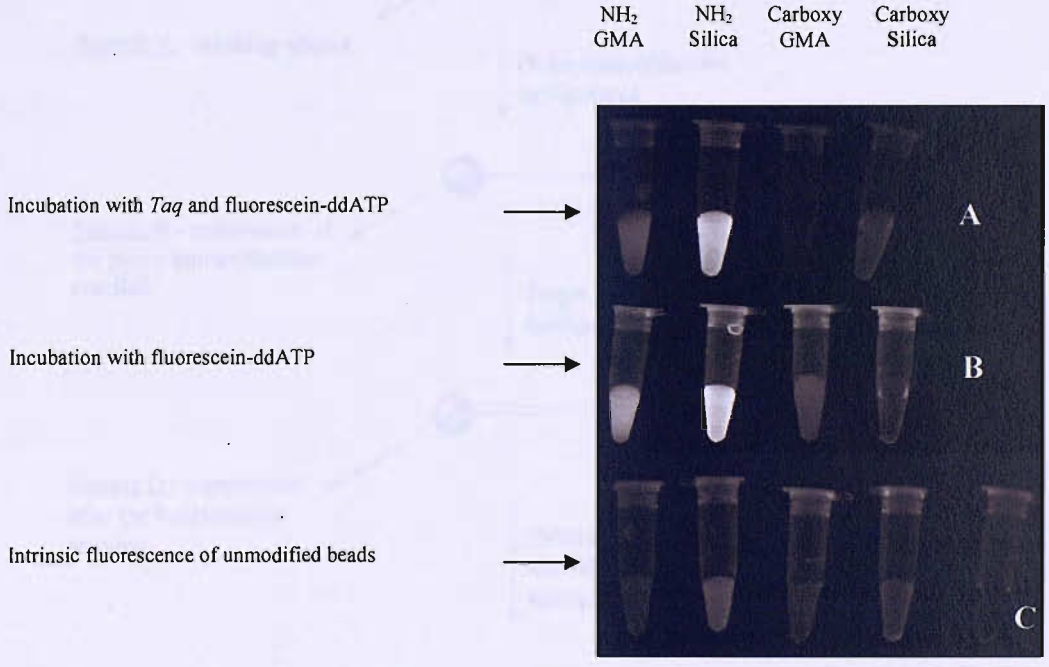


Figure 2.11 UV light results. (A) – Amino GMA, amino silica, carboxy GMA and carboxy silica – Incubation with enzyme and labelled nucleotide. (B) – Amino GMA, amino silica, carboxy GMA and carboxy silica – Incubation with labelled nucleotide but no enzyme. (C) – Amino GMA, amino silica, carboxy GMA and carboxy silica unmodified suspensions and a microtube containing PBS buffer (10 mM K_2HPO_4, 1.8 mM NaH_2PO_4, 0.15 M NaCl, pH 7.4).

2.3.5 Assessing efficiency of probe attachment and hybridisation

The probe immobilisation and hybridisation to target reactions had been thought to occur as the absorbance analysis at 260 nm of the supernatants of the reactions indicated a decrease in DNA left in solution. However, the uncertainty of the labelling process in the positive controls, as non-specific binding of the labelled nucleotide had been shown to occur, suggested also the possibility of non-specific binding of the DNA targets as opposed to real hybridisation.

The results and evidence gathered so far suggested the need to confirm and quantify each experimental step before moving on to the next one. In order to do that, the experiment outlined in Figure 2.12 was set up to monitor the progress of each experimental step through the analysis of aliquots.

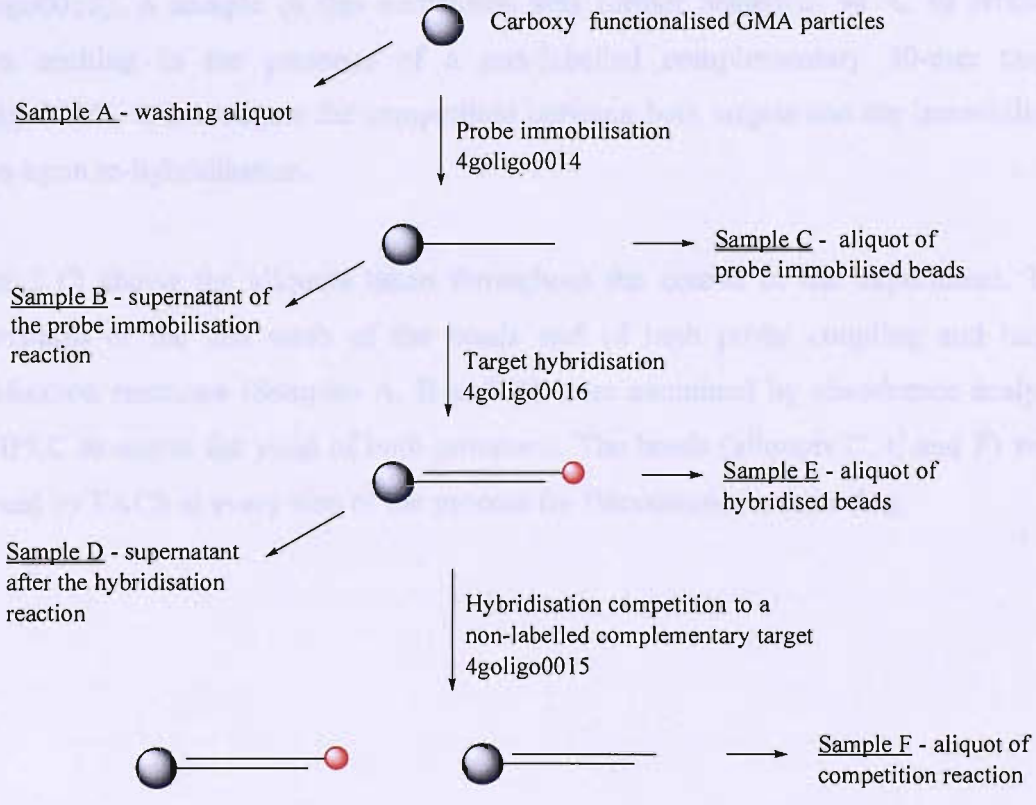


Figure 2.12 Experiment outline for the assessment of probe immobilisation and target hybridisation processes. Samples A, B and D were examined by absorbance and HPLC and aliquots C, E and F were analysed by FACS.

Oligonucleotide	Type of modification	Sequence 5' to 3'	Length
Probe – 4goligo0014	5' NH_2	GAGATGCACTCGAGTAAGTCAAGTCG	26
Target –4goligo0015	-	CCATCGACTTGACTTACTCGAGTGCATCTC	30
Target –4goligo0016	5' Cy5	CCATCGACTTGACTTACTCGAGTGCATCTC	30

Table 2.3 Oligonucleotides used for the assessment of probe immobilisation and efficiency of hybridisation.

Table 2.3 shows the oligonucleotide sequences used in the experiment. A 5' NH_2 26-mer probe (4goligo0014), was immobilised onto carboxy functionalised GMA beads, through amide chemistry in the presence of carbodiimide[32, 33] as a coupling agent. Carboxy GMA beads were chosen, due to their monodispersity and because they were less prone to non-specific binding of the labelled nucleotide (sections 2.3.1 and 2.3.4). The tethered probes were then incubated with a 5' Cy5 complementary 30-mer target, (4goligo0016). A sample of this suspension was further heated to 94 °C to promote duplex melting in the presence of a non-labelled complementary 30-mer target (4goligo0015), to investigate the competition between both targets and the immobilised probes upon re-hybridisation.

Figure 2.12 shows the aliquots taken throughout the course of the experiment. The supernatants of the last wash of the beads and of both probe coupling and target hybridisation reactions (Samples A, B and D) were examined by absorbance analysis and HPLC to assess the yield of both processes. The beads (aliquots C, E and F) were analysed by FACS at every step of the process for fluorescence monitoring.

Figure 2.13 shows the fluorescence of the beads at the initial stage of the experiment (C), after hybridisation of the immobilised probes to fluorescent target (E) and after the competition to the non-labelled target (F). Results show a substantial increase in Cy5 fluorescence after hybridisation. The signal decreases after the competition demonstrating that the non-labelled complementary strand is able to compete with the labelled target for the tethered probes.

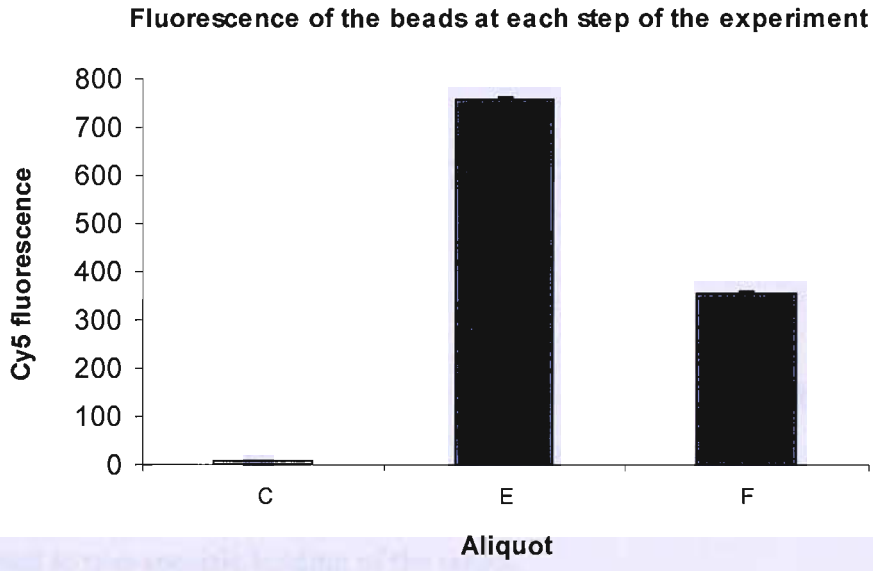


Figure 2.13 Fluorescence (± SE) of samples C, E and F at each step of the experiment. Fluorescence intensity increases when hybridisation to labelled target takes place and fluorescence drops when a non labelled complementary target competes with the labelled target for hybridisation to the tethered probes.

Table 2.4 shows the absorbance of samples A and B, at 260 nm. The absorbance of a stock solution of probe (4goligo0014) is included for reference. Results show the absence of absorption at 260nm in the washing before the coupling reaction, and shows that the amount of probe left in solution after the immobilisation step is minimal.

Sample	A_{260}
100 μM stock solution	1.837
A (before probe immobilisation)	0.007
B (after probe immobilisation)	0.072

Table 2.4 A_{260} of samples A and B and of a 100 μM stock solution of probe 4goligo0014 as reference.

Table 2.5 shows the absorbance of sample D at 646 nm. The absorbance of a stock solution of target (4goligo0016) is included for reference. Comparison between the absorbance of the target stock solution and the absorbance of the supernatant of the reaction after hybridisation permitted the calculation of the amount of unhybridised labelled target.

Sample	A_{646}
100 μM stock solution	1.593
D (after target hybridisation)	0.725

Table 2.5 A_{646} of sample D and of a 100 μM stock solution of target 4goligo0016 as reference.

The combined analysis of the absorbance spectra and HPLC (not shown) analysis for the samples examined at crucial steps of the experiment permitted the calculation of 98% probe coupling and 7% hybridisation yields assuming a correct duplex formation as opposed to non-specific binding of the target.

The unexpected low level of hybridisation regarding the high level of immobilised probe was surprising. This suggested the possibility of an elevated probe density resulting in a crowded environment not allowing the targets to access the probes. The use of spacers was considered as they are widely mentioned in the literature,[10, 11, 16, 30, 37] as a way to overcome this situation and improve the hybridisation yield. The utilisation of spacers is expected to result in the extension of the probe away from the solid support facilitating the hybridisation between probe and target.[19]

2.3.5.1 Single base chain extension in solution

Single base chain extension was attempted on immobilised probe 4goligo0014 hybridised to target, 4goligo0015, with Thermo Sequenase™ DNA polymerase. Table 2.6 shows the oligonucleotide sequences. Thermo Sequenase™ was chosen because of its excellent sequencing ability.[38] However, the fluorescence results obtained were similar to the ones observed before using the other DNA polymerases, also suggesting non-specific binding of fluorescein-ddATP (data not shown).

Oligonucleotide	Type of modification	Sequence 5' to 3'	Length
Probe – 4goligo0014	5' NH_2	GAGATGCACTCGAGTAAGTCAAGTCG	26
Target – 4goligo0015	-	CCATCGACTTGACTTACTCGAGTGCATCTC	30

Table 2.6 Oligonucleotides used for the assessment of probe immobilisation and efficiency of hybridisation.

In order to avoid this non-specific binding of the labelled nucleotide onto the solid support, a different approach was considered; single base chain extension in solution and subsequent immobilisation of the labelled duplex onto the beads. Single base chain extension was carried out in solution using Thermo Sequenase™. An incubation using *Taq* was also prepared for comparison. However, the immobilisation of the labelled duplexes onto GMA beads failed constantly and the beads never showed fluorescence.

To confirm chain extension was taking place in solution, non-denaturing PAGE (Polyacrylamide Gel Electrophoresis) analysis (Figure 2.14) of the incubation extensions with Thermo Sequenase™ and *Taq* were carried out. The analysis showed that single base chain extension was successfully achieved in solution using both polymerases. However, the extent of the chain extension carried out by *Taq* is minimal and most of the fluorescein-ddATP remains unattached (lanes 5 and 10) when compared to that of Thermo Sequenase™ (lanes 4 and 9). This is not surprising, as Thermo Sequenase™ incorporates ddNTPs much more efficiently than *Taq*.[38-40] Also, *Taq*

strongly discriminates against the incorporation of ddNTPs, especially if they bear a fluorescein label as is the case.[40-42]

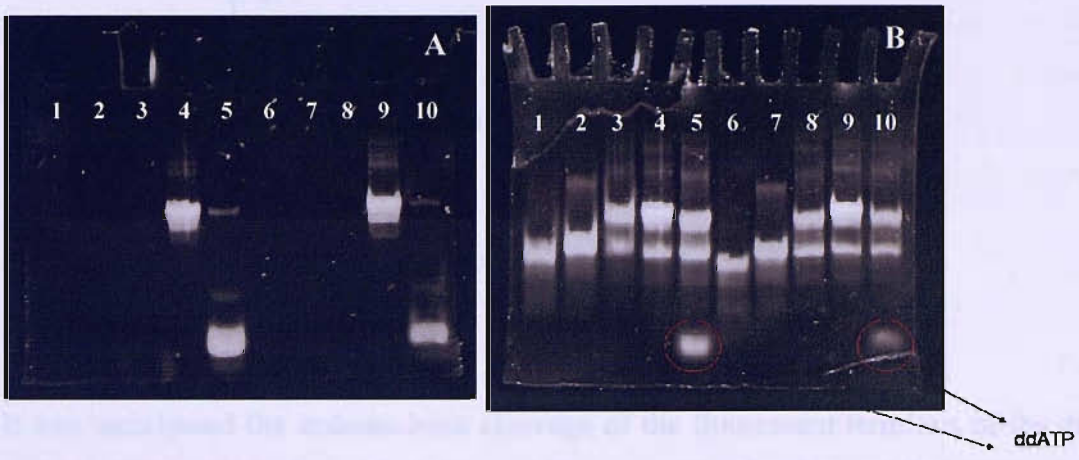


Figure 2.14 Non – denaturing PAGE fluorescence imaging: (A) unstained, showing the incorporation of a labelled nucleotide, (B) stained with SYBR® Green RNA; Lane 1 and 6 – Probe 4goligo0014; Lane 2 and 7 – Target 4goligo0015; Lane 3 and 8 – Hybridised Duplex; Lane 4 and 9 – *Thermosequenase* labelling; Lane 5 and 10 – *Taq* labelling.

The fact that chain extension occured in solution and not on the solid phase suggested steric hindrance from the solid support which inhibited the enzymes to work freely around the immobilised duplex. Also, the failed immobilisation of the labelled duplexes onto the beads suggested a steric impediment.

2.3.5.2 Restriction endonuclease digestion

To investigate further the possibility of steric hindrance from the solid support, an aliquot of immobilised probe 4goligo0014 hybridised to Cy5 labelled target 4goligo0016 was incubated with restriction enzyme *XhoI*. Probe 4goligo0014 and targets 4goligo0015/16 design included a restriction site recognised by *XhoI* formed upon hybridisation. Figure 2.15 shows the palindromic restriction site located in the middle of the double helix and where the enzyme cuts the DNA molecule, between nucleotides C and T in both directions generating two overhangs.

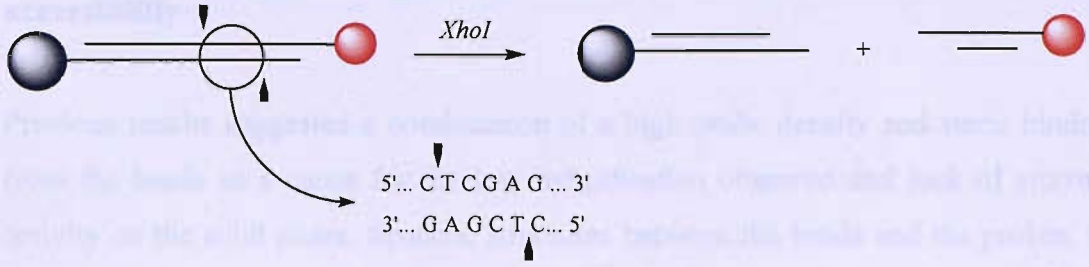


Figure 2.15 Restriction recognition site for *XhoI*.

It was anticipated the endonuclease cleavage of the fluorescent terminus of the duplex causing the beads to lose fluorescence. However, the incubation of the immobilised duplex with *XhoI* failed to produce significant digestion products (Figure 2.16) Results show that Cy5 fluorescence is practically the same before and after incubation with the restriction endonuclease, suggesting the enzyme did not digest the duplexes possibly due to the steric factor imposed on the construct by the solid phase.

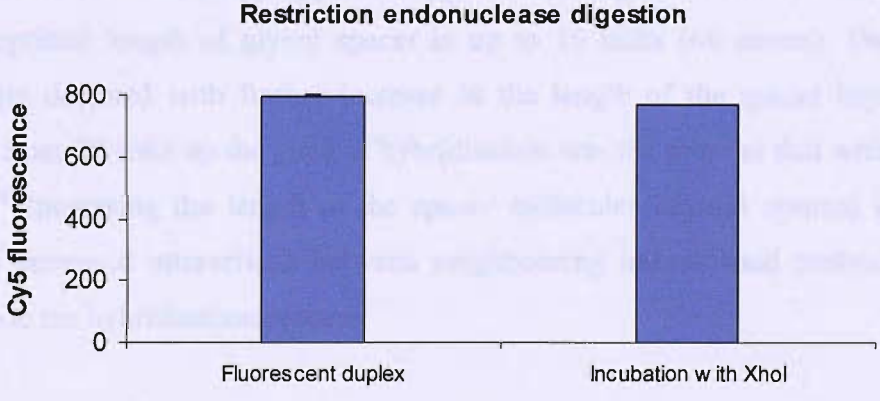


Figure 2.16 Fluorescence (± SE) of the beads before and after incubation with endonuclease.

2.3.6 Spacers as a mean to improve hybridisation efficiency and enzyme accessibility

Previous results suggested a combination of a high probe density and steric hindrance from the beads as a cause for the low hybridisation observed and lack of enzymatic activity on the solid phase. Spacers, structures between the beads and the probes, were explored in an attempt to improve hybridisation efficiency and enzyme accessibility on the immobilised duplexes. In theory, the more a coupled oligonucleotide probe is spatially removed from the solid support, the closer it is to solution state and the more likely it is to react freely to its target.[10, 30, 37]

The effect of spacers used between the solid supports and the DNA probes as a mean to enhance hybridisation is a controversial issue in the literature.[10, 16, 30, 43] Although spacers have been used to decrease the effects of solid supports, there has been no systematic study of the various factors affecting hybridisation behaviour. Properties of the spacer which are likely to be important include length, charge, hydrophobicity and solvation.[30] Whereas some authors [10, 37] observe that the use of spacers with different lengths has no noticeable effect on the hybridisation signal, other authors such as Southern et al. [30] reported a 150-fold increase in the hybridisation signal by inserting glycol spacers between polypropylene supports and oligonucleotides. Southern found that the optimal length of glycol spacer is up to 10 units (60 atoms). Duplex yield surprisingly declined with further increase in the length of the spacer beyond 10-12 units and from 30 units up the yield of hybridisation was the same as that with no spacer at all.[11, 30] Increasing the length of the spacer molecules beyond optimal length also results in increased interactions between neighbouring immobilised probes, which in turn impede the hybridisation process.

To investigate this matter and the implications of the use of spacers, a study involving spacers of various lengths was undertaken. GMA beads were modified with one, two and three amino-octanoic acid (Aoct) groups to produce different lengths of spacer, as shown in Figure 2.17A. The amount of free amino functional groups after each Aoct addition and Fmoc deprotection was determined by quantitative Fmoc and ninhydrin tests and it was calculated to be 6.7, 4.8 and 4.2 μmol per gram of beads, for beads

modified with one, two and three spacer units, respectively. The amino groups were then reacted with succinic anhydride resulting in terminal carboxy groups (Figure 2.17B).

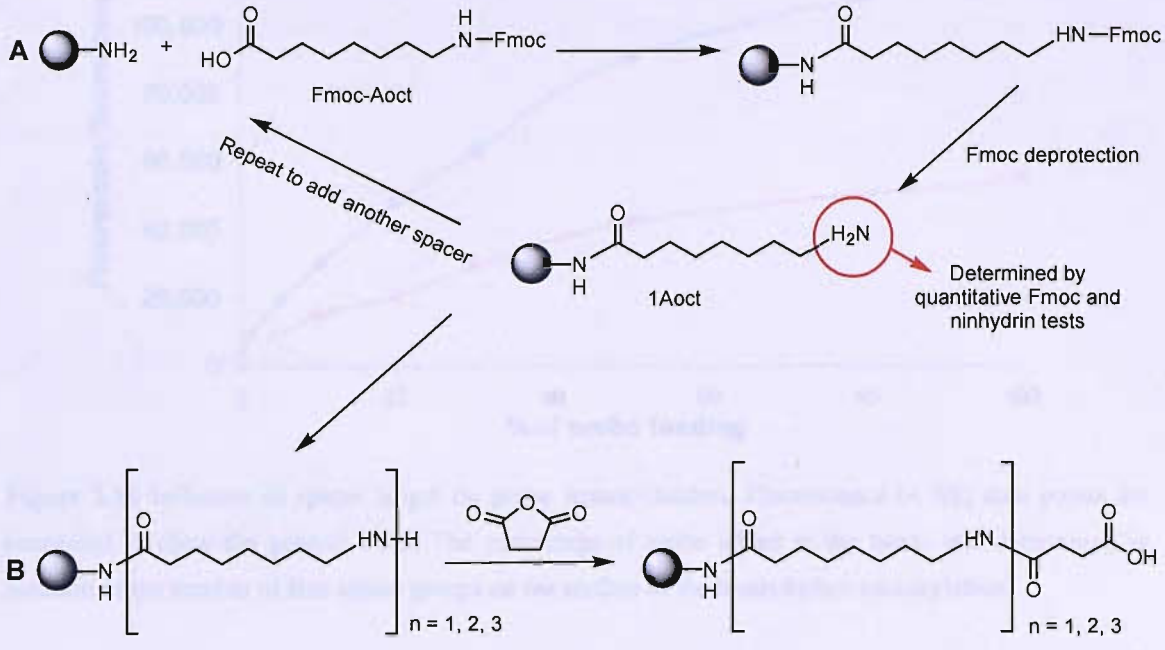


Figure 2.17 (A) GMA beads being modified with Aoct spacers to produce different lengths of spacer. After each Aoct addition and Fmoc deprotection the amount of free amino groups is determined by quantitative Fmoc and ninhydrin tests. (B) Conversion of the amino groups on GMA beads to carboxy groups by reaction with succinic anhydride.

Each batch of differently modified GMA beads was loaded with different percentages (ranging from 0 up to 100 %) of a 5' amino – 3' fluorescein 26-mer probe (4goligo0017) (Table 2.7). The percentage of probe added to the beads was determined as function of the number of free amino groups on the surface of the beads before succinylation. Fluorescence analysis of the beads is displayed in Figure 2.18.

Probe	Type of modification	Sequence 5' to 3'	Length
4goligo0017	5' NH_2 3' Fluorescein	GAGATGCACTCGAGTAAGTCAAGTCG	26

Table 2.7 Fluorescent probe used to assess the loading capacity of the beads.

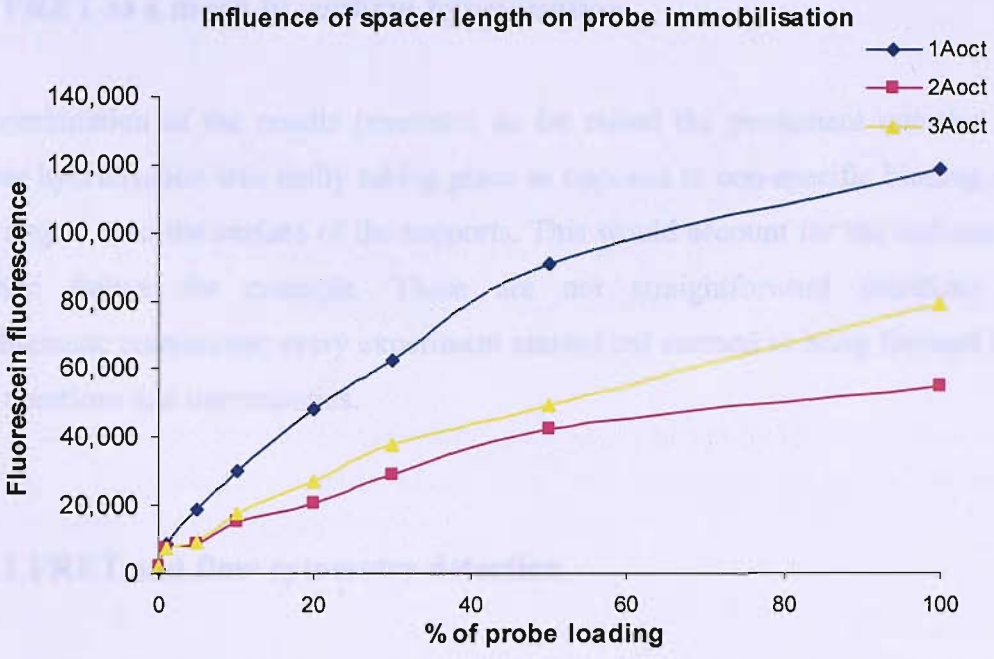


Figure 2.18 Influence of spacer length on probe immobilisation. Fluorescence (± SE) data points are connected to show the general trend. The percentage of probe added to the beads was determined as function of the number of free amino groups on the surface of the beads before succinylation.

The fluorescence trend exhibited in all cases is consistent with increasing loading of a fluorescent probe, tending towards a situation of fluorophore self-quenching, due to the extreme proximity of the probes at higher loadings. These results not only give information on the length of the spacer that provides the highest fluorescence signal, but also give information on avoiding high probe loading. Fluorescence signal decreased in the order 1Aoct › 3Aoct › 2Aoct. This can be explained by the consecutive chemistry steps taken for the addition of an additional spacer (leading to a decrease in the amount of free reactive groups on the beads). The possibility of formation of truncated species in the process must not be disregarded and may account for the lower fluorescence observed. Another explanation might be that increasing the length of the spacer increases the interactions between neighbouring immobilised probes, lowering the fluorescence due to self-quenching. GMA beads modified with 1Aoct spacer were subsequently used to pursue the studies on improving the hybridisation efficiency. Also, lower probe loadings, 1-20 %, were chosen in order to avoid a high probe density.

2.3.7 FRET as a mean to confirm hybridisation

The combination of the results presented so far raised the permanent question as to whether hybridisation was really taking place as opposed to non-specific binding of the DNA targets onto the surface of the supports. This would account for the endonuclease digestion failure for example. These are not straightforward situations with unproblematic resolutions; every experiment carried out seemed to bring forward a new set of questions and uncertainties.

2.3.7.1 FRET and flow cytometry detection

In order to confirm hybridisation, a FRET[44] construct outlined in Figure 2.19 involving a probe and a target modified with fluorescein and Cy5 fluorophores respectively was devised. [45]

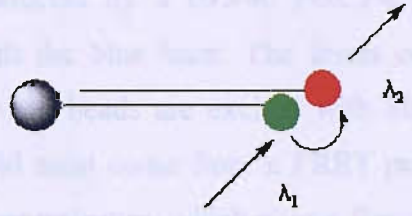


Figure 2.19 The hybridisation of fluorescently labelled probe and target generates FRET due to the close proximity of fluorescein and Cy5 fluorophores.

Figure 2.19 outlines the FRET process involving the radiationless transfer of energy from fluorescein to an appropriately positioned acceptor fluorophore, in this case Cy5. FRET occurs because the emission spectrum of fluorescein significantly overlaps (Figure 2.20) the excitation spectrum of Cy5. FRET happens provided that the donor and acceptor fluorophore dipoles are in a favourable mutual orientation and between 1-10 nm apart, which should be the case upon hybridisation.[46, 47]

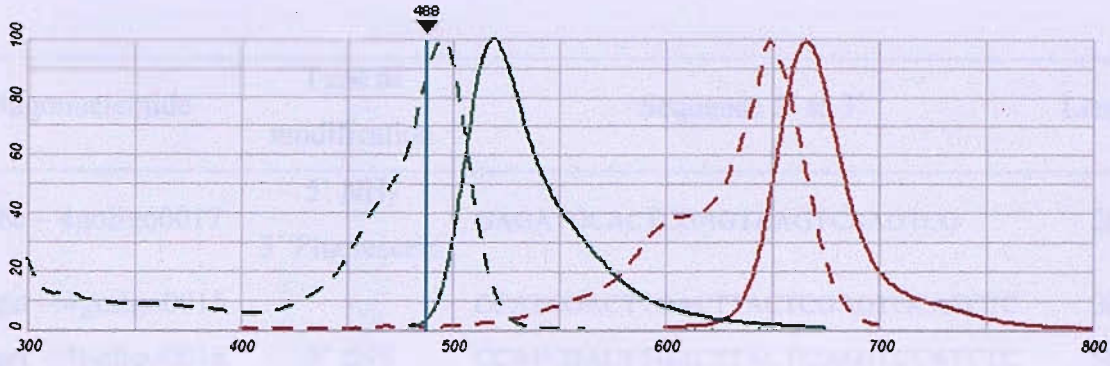


Figure 2.20 Fluorescein (Excitation λ= 494 – dotted green – and Emission λ= 520 – full green) and Cy5 (Excitation λ= 649 – dotted red – and Emission λ= 666 – full red) overlapping spectra.[48] A blue laser (488 nm) is used to excite fluorescein.

The FACS machine is configured with a blue (488 nm) and a red laser (633 nm). The blue laser excites fluorescein, which is detected by a 530/30 FITC-A band-pass filter and the red laser excites Cy5, which is detected by a 660/20 APC-A band-pass filter. FRET fluorescence is detected by a 695/40 PerCP-Cy5-5-A band-pass filter after fluorescein excitation with the blue laser. The lasers on the FACS machine have a configuration such that when beads are excited with blue laser, the emission signals recorded with filter 695/40 must come from a FRET process. The red laser is placed directly below the blue laser trajectory which allows fluorophores to relax after previous excitation explaining the Cy5 fluorescence detected by filter 660/20.

2.3.7.2 Confirmation of hybridisation

GMA beads modified with 1Aoct loaded with 1, 5 and 10 % of 5' amino – 3' fluorescein 26-mer probe (4goligo0017) respectively, underwent hybridisation to complementary targets, Cy5 labelled 4goligo0016 and to non-labelled 4goligo0015. Table 2.8 displays the oligonucleotides used in the FRET assay. Figure 2.21 shows the fluorescence of beads with increasing probe loading and Figures 2.22 and 2.23 show the fluorescence observed upon hybridisation to both targets.

Oligonucleotide	Type of modification	Sequence 5' to 3'	Length
Probe – 4goligo0017	5' NH_2 3' Fluorescein	GAGATGCACTCGAGTAAGTCAAGTCG	26
Target – 4goligo0015	-	CCATCGACTTGACTTACTCGAGTGCATCTC	30
Target – 4goligo0016	5' Cy5	CCATCGACTTGACTTACTCGAGTGCATCTC	30

Table 2.8 Oligonucleotides used for the assessment hybridisation efficiency through FRET.

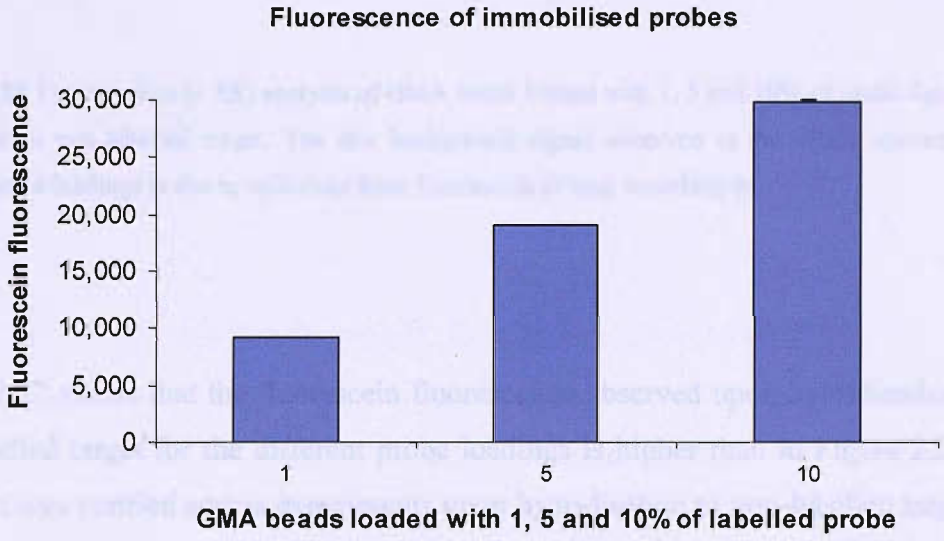


Figure 2.21 Fluorescence (± SE) of GMA beads with increasing probe loading. As explained before, the percentage of probe added to the beads was determined as function of the number of free amino groups on the surface of the beads before succinylation. A control of 1Aoct GMA beads (no probe) in the same experimental conditions yielded a fluorescence of 1,192±6.0.

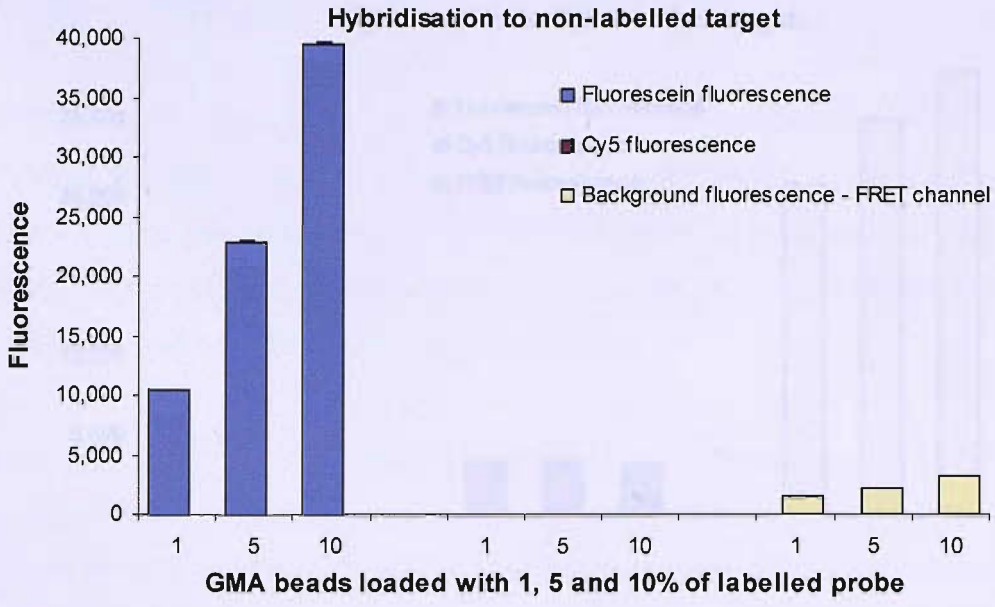


Figure 2.22 Fluorescence (± SE) analysis of GMA beads loaded with 1, 5 and 10% of probe 4goligo0017 hybridised to non-labelled target. The low background signal observed in the FRET channel for the various probe loadings is due to spill-over from fluorescein at long wavelengths.

Figure 2.22 shows that the fluorescein fluorescence observed upon hybridisation to the non-labelled target for the different probe loadings is higher than in Figure 2.21. This situation was verified across experiments upon hybridisation to non-labelled targets and may result from dequenching as single stranded DNA, due to its flexibility allows the fluorophore to interact with the DNA leading to fluorescence quenching.[49] Upon hybridisation, the rigidity of the duplex formed disrupts this quenching interaction and fluorescence increases.

As expected there is no Cy5 fluorescence and the low background signal observed in the FRET channel for the various probe loadings is due to spill-over from fluorescein at long wavelengths.

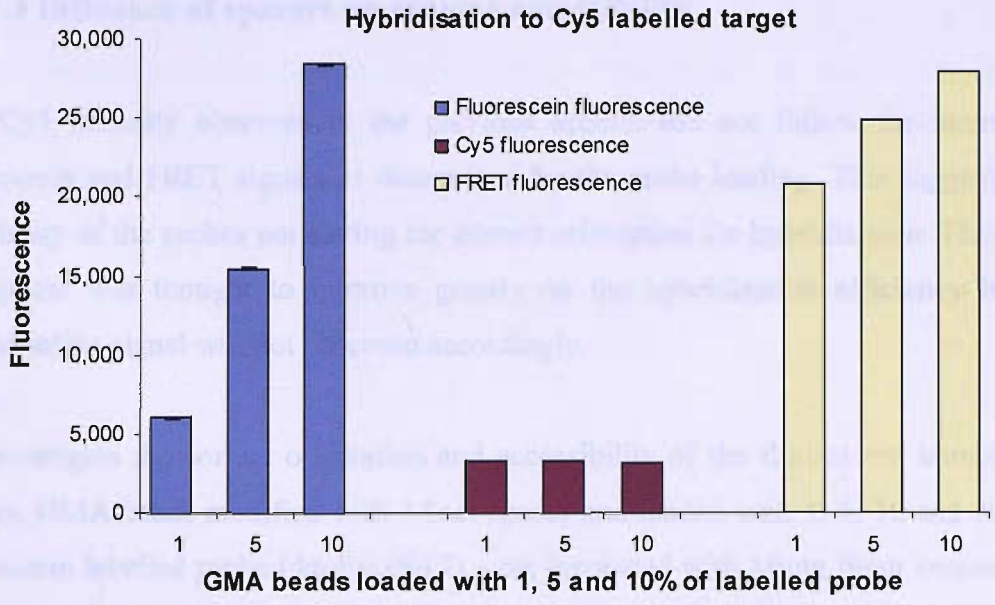


Figure 2.23 Fluorescence (± SE) analysis of GMA beads loaded with 1, 5 and 10% of probe 4goligo0017 hybridised to Cy5 labelled target. A control of 1Aoct GMA beads (no probe) incubated with Cy5 labelled target yielded Cy5 and FRET fluorescences of 925±4.6 and of 75±0.3, respectively.

Figure 2.23 shows that the fluorescein fluorescence upon hybridisation to the Cy5 labelled target is lower than in Figure 2.21. This was expected and can be explained by channelling of energy to Cy5, increasing the FRET signal. Results also show that the hybridisation to the Cy5 labelled target results in high Cy5 and very high FRET signals. The Cy5 fluorescence observed, with the use of a spacer, shows an improvement in hybridisation over the experiments carried out in previous sections. However, results are not clear as an increase in FRET was not accompanied by an increase in Cy5 fluorescence. This was not understood and suggested another effect taking part in the process.

FRET studies were furthered with the used of different FRET pairs; TAMRA and Cy5, Rox and Cy5, but the Cy5 and FRET fluorescences exhibited in all cases were not consistent with increasing probe loading. Also, the repeatability and reproducibility of results was not consistent across experiments, leading to the abandonment of these experiments.

2.3.7.3 Influence of spacers on enzyme accessibility

The Cy5 intensity observed in the previous section did not follow the increase in fluorescein and FRET signals as determined by the probe loading. This suggested the possibility of the probes not having the correct orientation for hybridisation. The use of the spacer was thought to improve greatly on the hybridisation efficiency but the hybridisation signal was not observed accordingly.

To investigate the correct orientation and accessibility of the fluorescent immobilised probes, GMA beads modified with 1Aoct spacer and loaded with 1, 5, 10 and 20 % of fluorescein labelled probe (4goligo0017) were incubated with Mung Bean exonuclease. This enzyme is single stranded DNA specific and degrades single stranded extensions from the ends of DNA. The cleavage of the 3' fluorescent end was expected to cause a decrease in fluorescence of the beads. Figure 2.24 shows the difference in fluorescence of the beads after incubation with the exonuclease

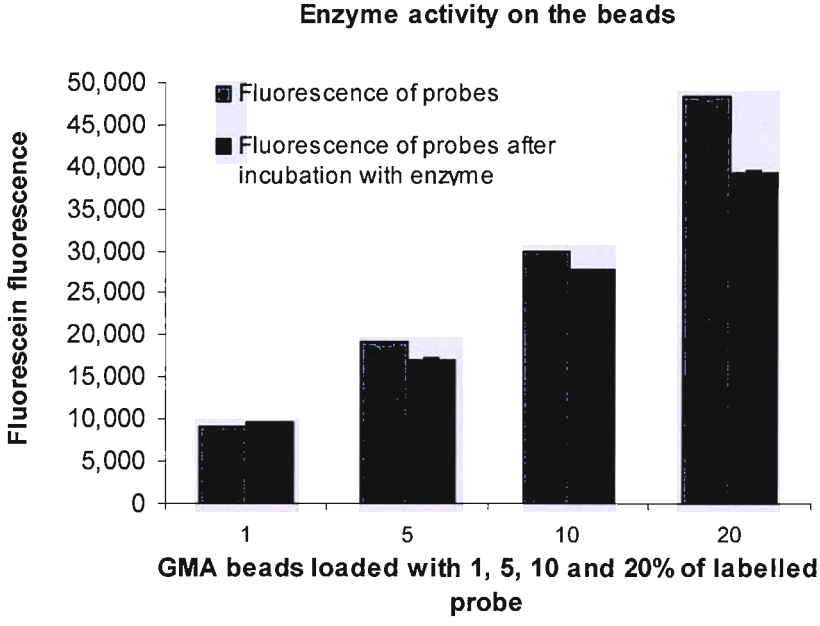
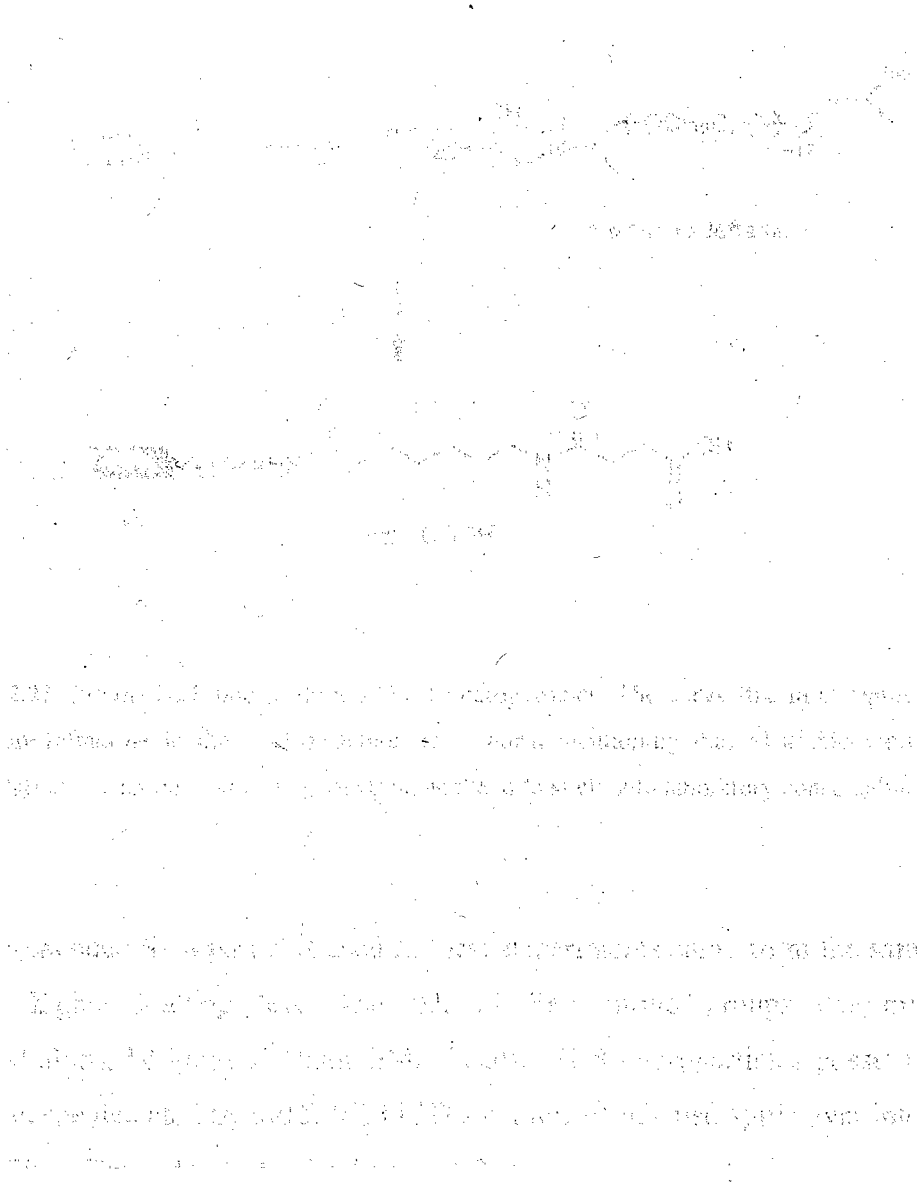


Figure 2.24 Fluorescence (± SE) of the beads after incubation with Mung Bean exonuclease.

Results showed that the decrease in fluorescence intensity of the beads after enzyme incubation, occurred only slightly, suggesting the exonuclease was unable to access and degrade the 3' termini of the majority of the probes. The combination of these results strongly suggested the probes were not correctly oriented and accessible for hybridisation and that might explain the poor level of hybridisation observed (low Cy5) and difficulty in enzyme accessibility to the duplex. However, it does not explain why the Cy5 signal observed did not follow the increase in fluorescein and FRET fluorescence.



2.3.8 Preliminary experiments on SU8

2.3.8.1 Influence of increasing probe loading on the fluorescence of SU8 microparticles

The probe immobilisation and target hybridisation processes were transferred to un-encoded SU8 microparticles, described in section 2.1.1. SU8 microparticles were modified with a linker *bis*-amino Jeffamine onto which an Fmoc-Aoct spacer was attached. Finally, the free amino groups obtained after Fmoc deprotection, as explained in section 2.3.6, were reacted with succinic anhydride resulting in terminal carboxy groups (Figure 2.25).

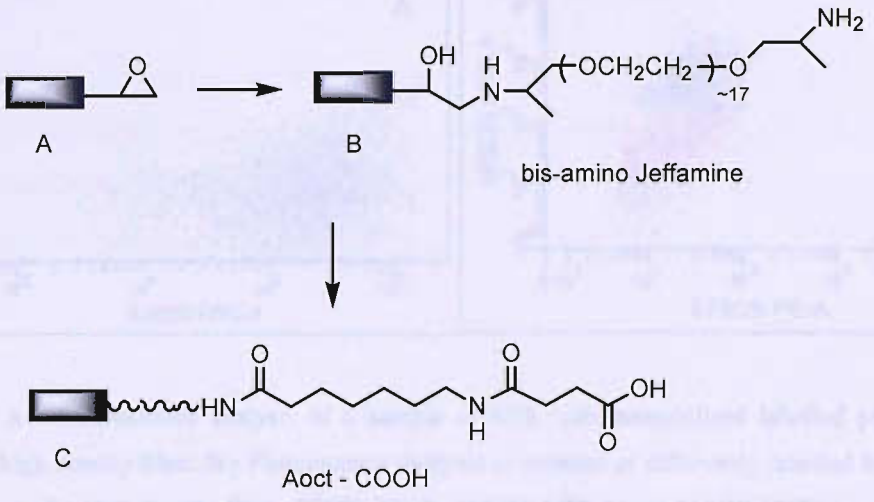


Figure 2.25 Summarised modification of SU8 microparticles. The curvy line in C represents the linker *bis*-amino Jeffamine in the final structure. It is worth mentioning that SU8 microparticles, prior to succinylation, are prone to severe aggregation and tend to stick onto laboratory consumables.

The un-encoded SU8 particles used in these experiments came from the same batch and had a higher loading level (number of free amino groups determined before succinylation), 18 μmol g^{-1}, than GMA beads. SU8 microparticles possess an intrinsic fluorescence detected by the 530/30 FITC-A filter which also spills over into the 575/26 PE-A filter. This adds to the background fluorescence when the attachment of labelled probes containing fluorophores detected by those filters is carried out. This problem is currently being addressed and taken into account at the fabrication process.

At the time of these experiments all samples of SU8 microparticles provided highly irreproducible scattered populations when analysed by FACS, such as the one depicted in Figure 2.26A. This was due to the three dimensional features of the microparticles which possessed different sizes. This made them difficult to study as a single centred population as in the GMA case. This problem has now been overcome due to the addition of a stronger neutral density filter on the FACS, introduced between the lasers and the flow of particles to be analysed. This has allowed other work in the group to distinguish two differently labelled populations in a multiplex assay[50] (Figure 2.26B).

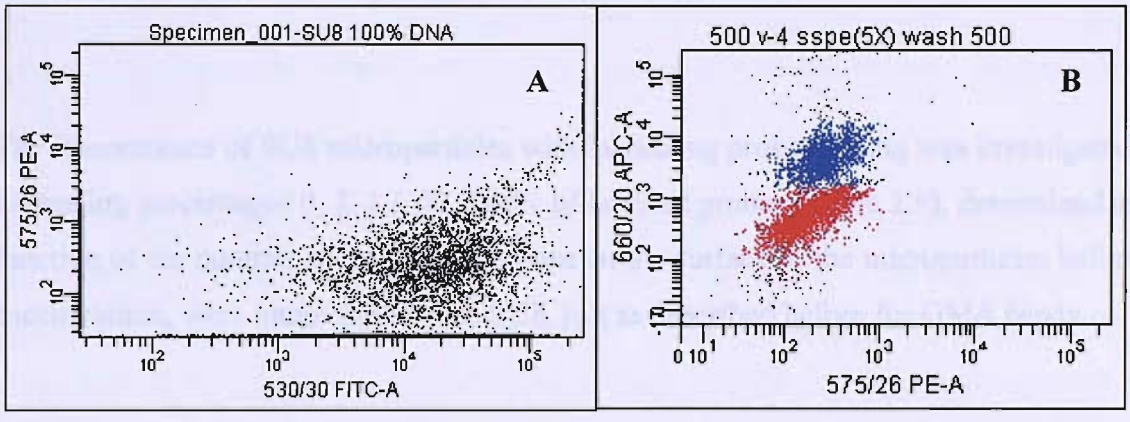


Figure 2.26 A - Fluorescence analysis of a sample of SU8 with immobilised labelled probe, in the absence of a high density filter. B - Fluorescence analysis of mixture of differently labelled SU8 samples in the presence of a high density filter. 575/26 PE-A, 660/20 APC-A and 530/30 FITC-A are the filters used to detect Cy3, Cy5 and fluorescein fluorescence, respectively.

As SU8 microparticles are heavier than GMA beads they settled out of solution very easily during the experiments and the analysis. The use of other solvents, as glucose, was explored, but the effects in population scattering worsened (Figure 2.27).

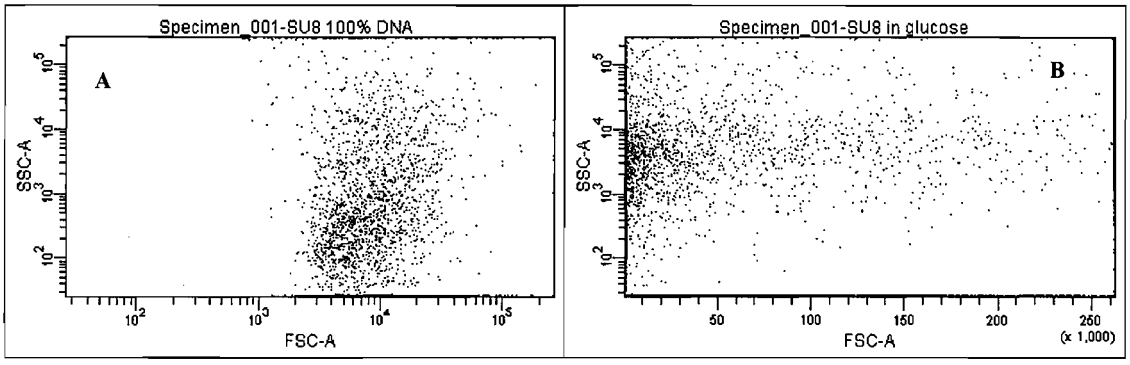


Figure 2.27 Scatter plots for a population of SU8 microparticles with immobilised labelled probes in aqueous media (A) and glucose (B). FSC-A (Forward scatter) and SSC-A (Side scatter) provide information about the size and morphology of the analysed microparticles.

The fluorescence of SU8 microparticles with increasing probe loading was investigated. Increasing percentages 0, 2, 10, 50, 100% of labelled probes (Table 2.9), determined as function of the number of free amino groups on the surface of the microparticles before succinylation, were immobilised onto SU8, just as described before for GMA beads.

Probes	Type of modification	Sequence 5' to 3'	Length
4goligo0017	5' NH_2 3' Fluorescein	GAGATGCACTCGAGTAAGTCAAGTCG	26
4goligo0030	5' Cy5 3' NH_2		

Table 2.9 Oligonucleotide probes used for assessing the influence of increasing probe loading on the fluorescence of SU8 microparticles.

Figure 2.28 shows that the increase in fluorescence observed is consistent with increasing probe loading for both fluorophores. However, the Cy5 labelled probe loading generated a higher signal than the fluorescein labelled probe loading.

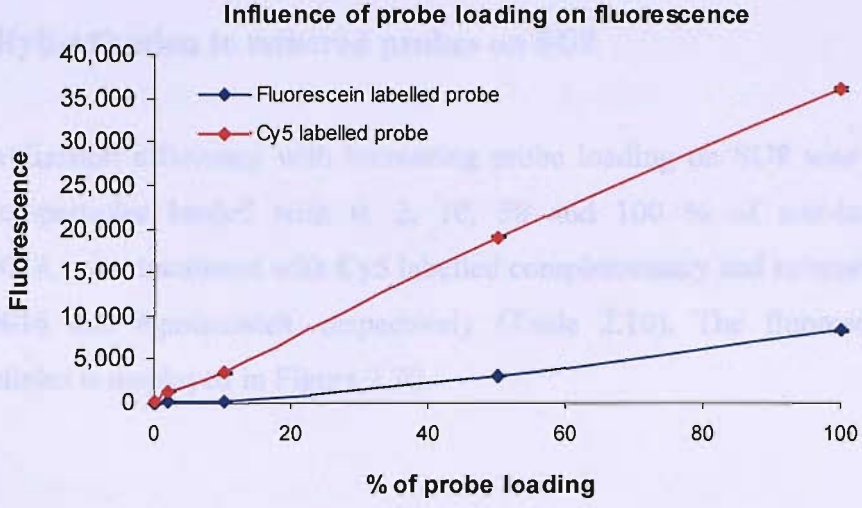


Figure 2.28 Influence of increasing probe loading on the fluorescence of SU8 microparticles. The percentage of probe added to the microparticles was determined as function of the number of free amino groups on the surface of SU8 before succinylation. Fluorescence (± SE) data points are connected to show the trend.

However, as also happened with GMA, the repeatability and reproducibility of the probe loading influence on the fluorescence was not consistent across experiments. As an example, Figure 2.29 shows the fluorescence observed with increasing Cy5 labelled probe loading for the same experiment done in triplicate. Although the observed trend is the same, the intensities exhibited are not. This was not understood as all experiments were carried out under the same conditions.

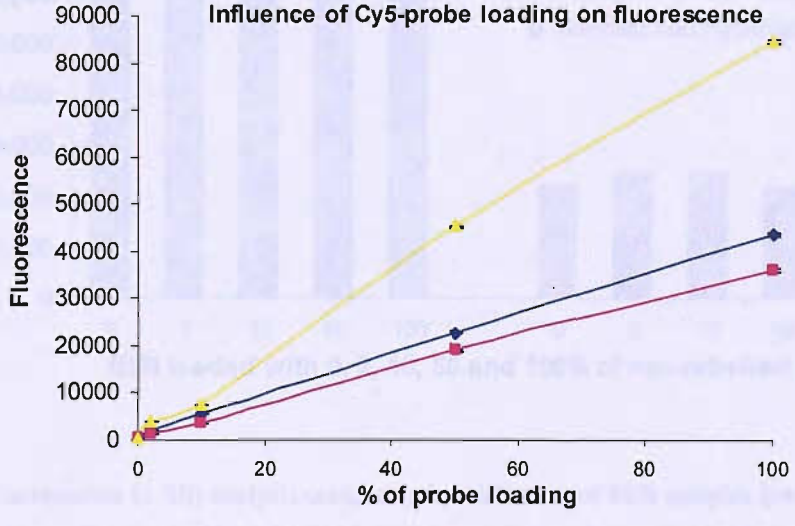


Figure 2.29 Reproducibility of Cy5 labelled probe loading on the fluorescence of SU8 microparticles; the same experiment done in triplicate generated three different results (coloured lines). Fluorescence (± SE) data points are connected to show the trend.

2.3.8.2 Hybridisation to tethered probes on SU8

The hybridisation efficiency with increasing probe loading on SU8 was investigated. SU8 microparticles loaded with 0, 2, 10, 50 and 100 % of non-labelled probe 4goligo0014, were incubated with Cy5 labelled complementary and mismatched targets, 4goligo0016 and 4gmismatch, respectively (Table 2.10). The fluorescence of the microparticles is displayed in Figure 2.30.

Oligonucleotide	Type of modification	Sequence 5' to 3'	Length
Probe – 4goligo0014	5' NH_2	GAGATGCACTCGAGTAAGTCAAGTCG	26
Target – 4goligo0016	5' Cy5	CCATCGACTTGACTTACTCGAGTGCATCTC	30
Target – 4gmismatch	5' Cy5	CGAAGTCACGATGTCCACTGTTC	23

Table 2.10 Oligonucleotides used for the assessment of efficiency of hybridisation on SU8.

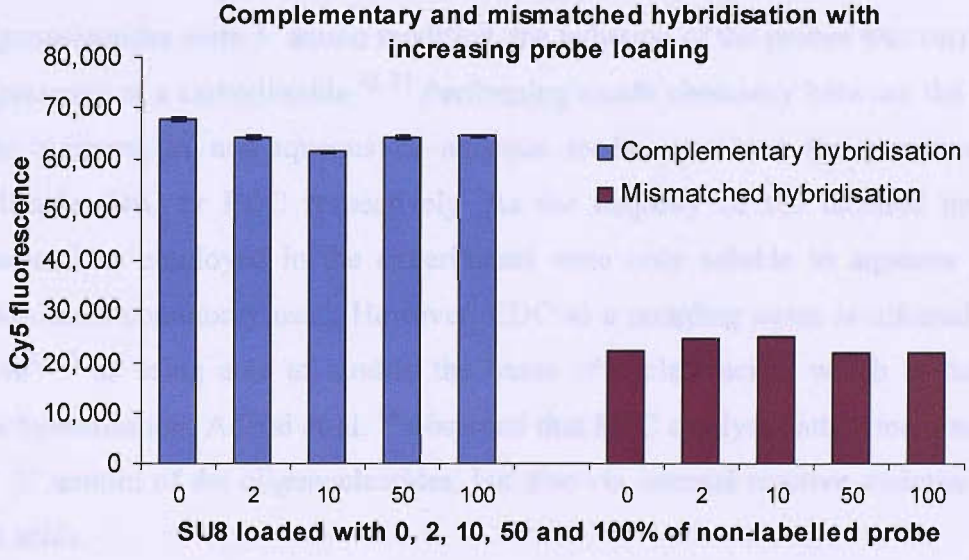


Figure 2.30 Fluorescence (± SE) analysis comparing hybridisation of SU8 samples loaded with different percentages of probe 4goligo0014 to labelled complementary and mismatched targets. As previously explained, the percentage of probe added to the microparticles was determined as function of the number of free amino groups on the surface of SU8 before succinylation. An SU8 unmodified control yielded a Cy5 fluorescence of 345±1.7.

Figure 2.30 shows that the level of fluorescence in the absence of probe (0 %) for both complementary and mismatched incubations exhibits approximately the same signal as the other samples. This clearly suggests non-specific binding of the labelled targets onto the support as opposed to hybridisation. As the Cy5 intensity is higher for the complementary target than it is for the mismatch it is assumed the oligonucleotides have different non-specific bindings.

2.3.9 Chemistry of probe immobilisation

Some authors[51, 52] are in agreement that the accessibility of immobilised DNA depends more significantly on the process of immobilisation rather than on the type of support used for fixation. When the amino functionalised beads, used in the preliminary single base chain extension experiments, were proved to exhibit high levels of non-specific binding of the labelled nucleotide, their use was abandoned. Carboxy modification of amino functionalised beads was then carried out as the carboxy function appeared to be less prone to that non-specific binding. But the change in functional groups on the surface of the beads implied a modification of the process of probe immobilisation. As the oligonucleotides were 5' amino modified, the tethering of the probes was carried out in the presence of a carbodiimide.[32, 33] Performing amide chemistry between the probes and the supports in non-aqueous or aqueous media, involved the presence of a carbodiimide, DIC or EDC respectively. As the majority of the labelled modified oligonucleotides employed in the experiments were only soluble in aqueous media, EDC was most commonly used. However, EDC as a coupling agent is referred in the literature[6, 20] as being able to modify the bases of nucleic acids, which in turn will impede hybridisation. Adessi et al. [16] observed that EDC catalyses attachment not only via the 5' termini of the oligonucleotides, but also via internal reactive moieties of the nucleic acids.

2.3.9.1 Investigating the EDC mediated coupling chemistry

The influence of EDC on the probe attachment was investigated in an attempt to explain some of the results gathered so far, such as the low levels of hybridisation and the fact that the hybridisation fluorescence did not follow the increase in probe loading.

A 5' NH_2 - 3' fluorescein probe, 4goligo0025, was loaded in different percentages, 0, 2, 10, 50, 100, and 200 % onto carboxy 1Aoct GMA beads using EDC mediated chemistry. A second experiment was set up without the presence of EDC to evaluate the amount of non-specific binding of the probe with its increasing concentration. A third experimental set, contained EDC and a 3' fluorescein probe, 4goligo0019 (no 5' modification) and aimed to investigate the amount of linkage through the internal reactive moieties of the nucleic acids. Table 2.11 shows the nature of the probes utilised on the experiment.

Probes	Type of modification	Sequence 5' to 3'	Length
4goligo0025	5' NH_2 3' Fluorescein	TTTTTTTTTGAGATGCACTCGAGTAAGTCAAGTCG	35
4 goligo0019	3' Fluorescein	GAGATGCACTCGAGTAAGTCAAGTCG	26

Table 2.11 Oligonucleotides used for the investigation of the role of EDC in the probe immobilisation.

According to Figure 2.31, probe 4goligo0019 (yellow) displays more than half of the fluorescence exhibited by probe 4goligo0025 (blue). As probe 4goligo0019 has no 5' amino modification, results imply that immobilisation occurred via the internal moieties of the oligonucleotide. There is also a substantial non-specific binding of probe 4goligo0025 onto the support as evidenced by the pink trend.

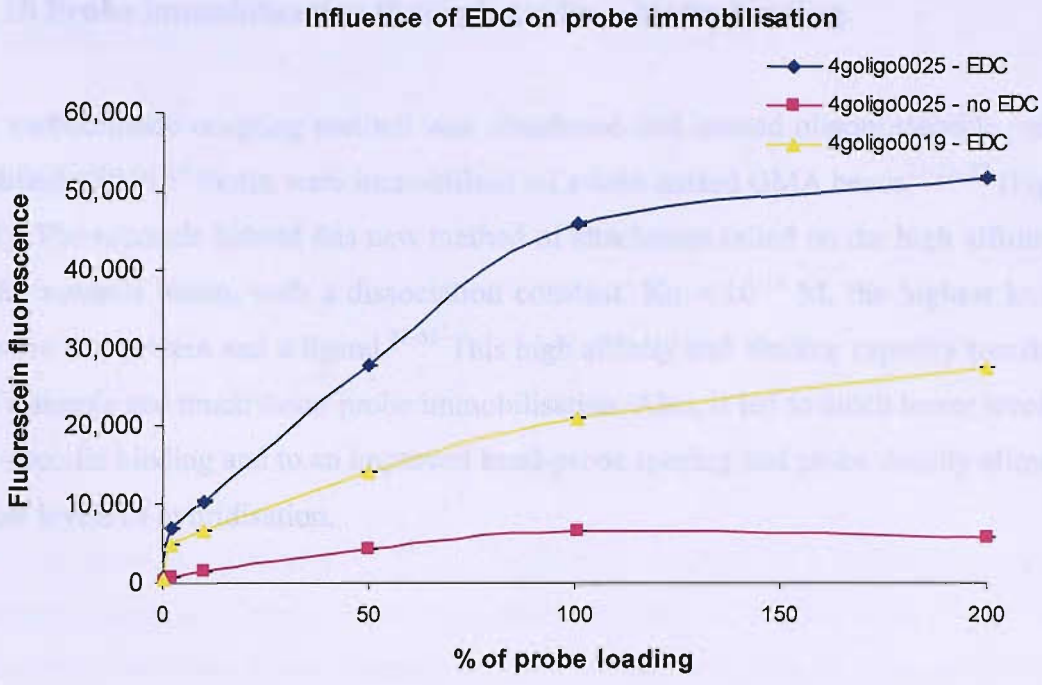


Figure 2.31 Assessing the influence of EDC on probe immobilisation. The percentage of probe added to the beads was determined as function of the number of free amino groups on the surface of the beads before succinylation. Fluorescence (± SE) data points are connected to show the trend.

These results were consistent in trend and in fluorescence intensity across experiments, verifying the unreliable action of EDC as a coupling agent. As the amount of probe added to the beads increases, so does the extent of incorrect immobilisation. This suggested the probes were not in a correct orientation and accessibility for target hybridisation and may explain the low levels of duplex formation observed on GMA. It also may explain why the hybridisation signal did not follow the increase in probe levels seen on both GMA and SU8 and may explain the poor or non-existent enzyme accessibility.

2.3.10 Probe immobilisation through avidin – biotin binding

The carbodiimide coupling method was abandoned and instead oligonucleotide probes modified with a 5' biotin were immobilised on avidin coated GMA beads[21-23, 27] (Figure 2.32). The rationale behind this new method of attachment relied on the high affinity of avidin towards biotin, with a dissociation constant, $K_D \approx 10^{-15}$ M, the highest known between any protein and a ligand.[53-55] This high affinity and binding capacity translated into a simple and much faster probe immobilisation. Also, it led to much lower levels of non-specific binding and to an improved bead-probe spacing and probe density allowing higher levels of hybridisation.

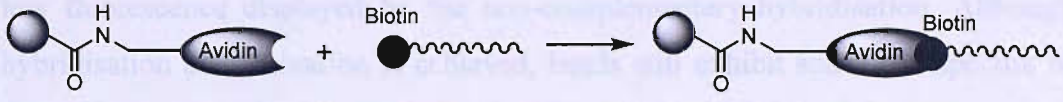


Figure 2.32 Avidin molecules were immobilised onto carboxy GMA beads via their amino lysine residues. Although avidin is a tetrameric protein attaching four molecules of biotin per tetramer, only one attachment is shown for simplicity of scheme.

5' biotinylated 16-mer probes, ATD0748 and ATD0747, were immobilised on avidin coated GMA beads. Each probe was then hybridised to its respective 5' Cy5 complementary target, ATD0750 and ATD0749. The level of non-complementary hybridisation of each probe was assessed by incubation with a 5' Cy5 non-complementary target, 4goligo0016. The hybridisation methodology was also revised and modified;[56-59] a high ionic strength buffer was used, suspensions were incubated overnight with constant agitation[57, 59] at constant temperature, and increased post-hybridisation stringency was utilised in the washing of the beads. Table 2.12 shows the oligonucleotides sequences employed in the experiment.

Oligonucleotides	Type of modification	Sequence 5' to 3'	Length
Probe - ATD0748	5' biotin	CTAGTTACTCTTGTTC	16
Probe - ATD0747	5' biotin	TTGTTATAGTTCTCTC	16
Target - ATD0750	5' Cy5	GAACAAGAGTAACTAG	16
Target - ATD0749	5' Cy5	GAGAGAACTATAACAA	16
Target - 4goligo0016	5' Cy5	CCATCGACTTGACTTACTCGAGTGCATCTC	30

Table 2.12 – Oligonucleotides used for the avidin – biotin probe immobilisation method.

After overnight incubation at 45 °C, the beads were washed with buffers of increased stringency and fluorescence was measured. Figure 2.33 shows how the high fluorescence exhibited for complementary hybridisation in both cases contrasts with the low fluorescence displayed by the non-complementary hybridisation. Although high hybridisation discrimination is achieved, beads still exhibit some non-specific binding from the non-complementary target, as comparison with the controls suggested (Figure 2.33).

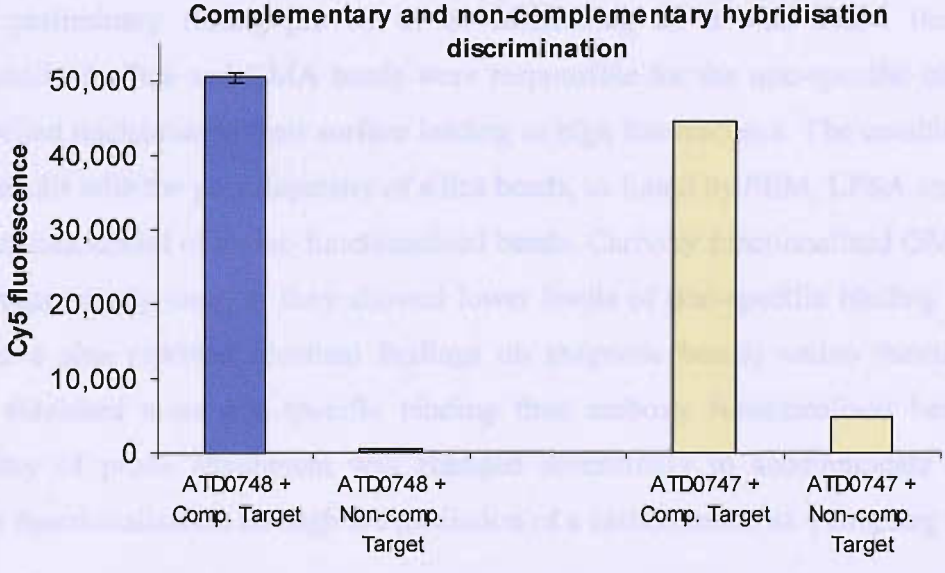


Figure 2.33 Fluorescence (± SE) of the beads after being washed with buffers: 5xSSPE (0.75 M NaCl, 50 mM NaH_2PO_4, 5 mM EDTA, pH 7.4), 0.1% SDS, 2xSSC (0.3 M NaCl, 30 mM Na_3citrate, pH 7.0), 0.1% SDS and 1xSSC (0.15 M NaCl, 15 mM Na_3citrate, pH 7.0), 0.1% SDS. Controls of unmodified GMA beads and avidin coated GMA incubated with target ATD750, yielded Cy5 fluorescences of 62±11 and 347±8, respectively.

2.4 Conclusion

This chapter describes the introductory approach to detecting DNA on the solid phase by reproducing the literature. However, the replication of the published experimental techniques was not straightforward to implement and the results obtained from every experiment would bring forward new issues.

GMA (5 μm), silica (5 and 10 μm) beads and SU8 microparticles were chosen as the solid supports to investigate on-bead DNA chemistry. The preliminary SBCE reactions involved amino modified DNA probes immobilised on the surface of amino GMA and silica beads through glutaraldehyde mediated chemistry.[32] The immobilised probes served as primers in a polymerase catalised extension reaction in which a fluorescently labelled ddATP, complementary to the next base in the target was intended to be incorporated in the structure. A decrease in absorbance at 260 nm in the supernatants of probe immobilisation and hybridisation to target reactions suggested the confirmation of the processes. Highly fluorescent beads seemed to suggest base extension had taken place.

These preliminary results proved to be misleading as it was found that amino functionalised silica and GMA beads were responsible for the non-specific binding of the labelled nucleotide to their surface leading to high fluorescence. The combination of these results with the polydispersity of silica beads, as found by SEM, LPSA and FACS, led to abandonment of amino functionalised beads. Carboxy functionalised GMA beads were subsequently used, as they showed lower levels of non-specific binding. Lund et *al.*[20] have also reported identical findings on magnetic beads; amino functionalised beads exhibited more non-specific binding than carboxy functionalised beads. The chemistry of probe attachment was changed accordingly to accommodate the new surface functionalisation through the mediation of a carbodiimide as a coupling agent.[32]

A quantitative protocol was set up to assess the efficiency of probe attachment and target hybridisation. Results of 98% and 7% were obtained for the efficiency of both procedures respectively. The low level of hybridisation in comparison to the high level of immobilised probe suggested an elevated probe density not allowing the targets to

access probes. The unsuccessful SBCE and restriction digestions carried out on the constructs suggested steric hindrance from the solid support not allowing the enzymes to access the duplexes. The combination of these two factors led to the use of spacers, structures placed between the solid support and the probe, to enable a solution-like environment for the probes and easy access for DNA targets and enzyme activity. However, the results obtained in subsequent experiments did not show an improvement on hybridisation or enzymatic activity. FRET was utilised as a way to confirm hybridisation, but the lack of reproducibility and repeatability of results led to the abandonment of the experiments. The lack of significant exonuclease digestion on fluorescent single-stranded immobilised probes, suggested the probes were not correctly orientated and readily accessible.

The chemistry of probe attachment was then investigated. Experimental results showed that carbodiimide driven immobilisation of probes, led to more than half of them being tethered through the side chain moieties.[16] This suggested an incorrect probe orientation resulting in a limited accessibility for target hybridisation.[6] This may explain the low levels of duplex formation observed on GMA, FRET irreproducibility and poor enzyme accessibility.

The construct avidin-biotin was then used to replace the carbodiimide mediated attachment. GMA beads were functionalised with avidin and DNA probes were modified with 5' biotinylated termini. The combination of a new tethering approach and a revised hybridisation methodology provided high levels of fluorescence and hybridisation discrimination exhibited. The new hybridisation methodology involved a high ionic strength hybridisation buffer, the continuous agitation of the immobilised probes and the targets in solution,[57, 59] and an increased post-hybridisation stringency washing procedure.[58] Continuous agitation as opposed to the occasional agitation carried out preliminarily, seemed to be more efficient in overcoming the diffusion problems encountered by the target in solution and also in allowing the beads (probes) to be in constant motion and to be accessible. The post-hybridisation stringency procedure enabled discrimination between complementary and non-complementary hybridisation reducing greatly the non-specific binding of the target.

SU8 un-encoded particles, developed within the 4G sequencing project were also utilised for these assays. Initially, these particles exhibited great tendency to aggregation and their three-dimensional features originated some difficulty in flow cytometry analysis. These situations have now been respectively resolved, with the used of buffers containing trace amounts of detergent, typically 0.02 % Tween and the introduction of a higher density filter on the FACS. The experiments carried out on un-encoded SU8 showed that the increase in fluorescence exhibited by the particles is consistent with the increase in fluorescent probe loading. However, the Cy5 fluorescence observed upon hybridisation did not follow the same trend. In addition, unmodified SU8 microparticles incubated with fluorescent target exhibited the same level of fluorescence as the microparticles loaded with probe, suggesting non-specific binding of the target as opposed to hybridisation. Ongoing optimisation experiments carried out by the group on encoded microparticles have also utilised the avidin-biotin construct for probe immobilisation, and results provided have shown a great improvement over the previous attachment protocol. Also, immunoassays and hybridisation discrimination studies carried out in a multiplex manner showed promising preliminary results for the implementation of the encoded particles.

The experimental results described here highlight the concern and attention probe immobilisation procedures deserve as they may severely compromise DNA solid phase detection methods. The immobilisation procedure has been shown to impart influence on the extension of hybridisation and enzyme accessibility to the duplexes. Also, the nature of the solid supports and the nature of the functional reactive groups, used to tether the probes might exert influence on the level of non-specific binding generating misleading results.

The promising results obtained using the avidin-biotin probe attachment and the revised hybridisation protocol enabled the experiments on hybridisation selectivity carried out in chapter 3 and a detailed study on hybridisation discrimination to single point mutation carried out in chapter 4. Also, SBCE and restriction endonuclease digestion were successfully made to work on the immobilised duplexes permitting the solid phase enzyme accessibility studies in chapter 5.

2.5 Materials and Methods

2.5.1 Reagents

Carboxylate-modified silica beads (5 - 50 mg/ml) and Amino-modified silica beads (5 μm - 50 mg/ml) were obtained from GbR Kisker. Amino-modified silica beads (10 μm) were purchased from Thermo Hipersil. Amino-modified GMA beads (5 - 6 μm - 10% (W/V)) were obtained from Bangs Labs. Un-encoded SU8 microparticles were fabricated within the research group. Oligonucleotide sequences were synthesised by ATDbio and Sigma-Aldrich (Tables 2.13 and 2.14). Avidin DN (1.0 mg/ml) was from Vector Laboratories. *Taq* DNA polymerase provided with 10x PCR buffer and 25 mM $MgCl_2$ solution, was purchased from Sigma-Aldrich. *Pfu Turbo* DNA polymerase provided with 10x cloned *Pfu* buffer was obtained from Stratagene. Thermo Sequenase™ DNA polymerase (with *Thermoplasma acidophilum* Inorganic Pyrophosphatase), provided with 1x Thermo Sequenase reaction buffer and 1x Thermo Sequenase dilution buffer, was obtained from Amersham Biosciences. *XhoI* provided with 10x NEB 2 buffer came from New England Bio Labs Inc. Mung Bean Nuclease, provided with 10x Mung Bean Nuclease reaction buffer came from New England Bio Labs Inc. Fluorescein – 12 – ddATP was purchased from PerkinElmer Life Sciences. 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC), 99%, Succinic Anhydride, 99%, N,N – Dimethylformamide (DMF), 99%, N,N' Diisopropyl - carbodiimide (DIC), 99%, 1 – Hydroxybenzotriazole (HOBt), 99%, N,N' Diisopropylethilamine (DIPEA), 99%, and Acetic Anhydride, 99%, came from Acros. Glutaraldehyde, 50%, Acetonitrile and THF Laboratory Reagent Grade were obtained from Fisher Scientific. Triethylamine (TEA), 99%, was purchased from Avocado. 4 – Dimethylaminopiridine (DMAP) came from Fisher BioReagents. FMOC–8–Aoct-OH (Fmoc amino-octanoic acid), › 99%, came from Fluka. Imidazole, 99.00% + came from Aldrich. All other reagents required for the preparation of coupling, storage and post-hybridisation stringency buffers; potassium phosphate dibasic, potassium phosphate monobasic, boric acid, sodium tetraborate, sodium chloride, sodium citrate, sodium dihydrogen phosphate, tris-HCl, EDTA, SDS (sodium dodecyl sulphate), hydroxylamine were of analytical grade and were obtained from Acros.

DNA Probes	Type of modification	Sequence 5' – 3'	Length (nt)
4goligo0006	5' NH_2	GAGATGCCCAGAATGAAGTCAAGTCG	26
4goligo0014	5' NH_2	GAGATGCACTCGAGTAAGTCAAGTCG	26
4goligo0017	5' NH_2 3' Fluorescein	GAGATGCACTCGAGTAAGTCAAGTCG	26
4goligo0030	5' Cy5 3' NH_2	GAGATGCACTCGAGTAAGTCAAGTCG	26
4goligo0019	3' Fluorescein	GAGATGCACTCGAGTAAGTCAAGTCG	26
4goligo0025	5' NH_2 3' Fluorescein	TTTTTTTTTGAGATGCACTCGAGTAAGTCAAGTCG	35
ATD0748	5' biotin	CTAGTTACTCTTGTTC	16
ATD0747	5' biotin	TTGTTATAGTTCTCTC	16

Table 2.13 DNA probe sequences employed.

DNA Targets	Type of modification	Sequence 5' – 3'	Length (nt)
4goligo0004	-	CCATCGACTTGACTTCATTCTGGGCATCTC	30
4goligo0011	-	CCATCGTCTTGACTTCATTCTGGGCATCTC	30
4goligo0015	-	CCATCGACTTGACTTACTCGAGTGCATCTC	30
4goligo0016	5' Cy5	CCATCGACTTGACTTACTCGAGTGCATCTC	30
4gmismatch	5' Cy5	CGAAGTCACGATGTCCACTGTTC	23
ATD0750	5' Cy5	GAACAAGAGTAACTAG	16
ATD0749	5' Cy5	GAGAGAACTATAACAA	16

Table 2.14 DNA target sequences employed.

2.5.2 FACS analysis

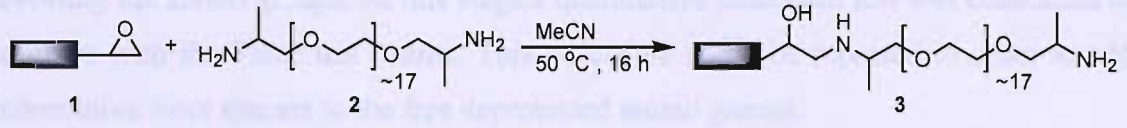
FACS analysis was performed on a BD Biosciences, FACS Aria. The excitation optics involved a Coherent® Sappire™ Solid State (488 nm) and SDS Uniphase™ HeNe air cooled (633 nm) lasers and fluorescence was detected with 530/30-A FITC, 660/20 APC-A, 575/26 PE-A and 695/40 PerCP-Cy5-5-A Band-pass filters and analysed using BD FACSDiva™ software. The laser delay was calibrated for all experiments and the PMT voltages remained unchanged throughout the experiments.

2.5.3 SEM and LDPS analysis

Silica (5 and 10 μm) and GMA (5 μm) beads were characterised by Scanning Electron Microscopy (SEM) on a JEOL JSM-5910. Silica (5 and 10 μm) were analysed on a Beckman-Coulter™ - Laser Diffraction Particle Size Analyser.

2.5.4 Functionalisation of microparticles

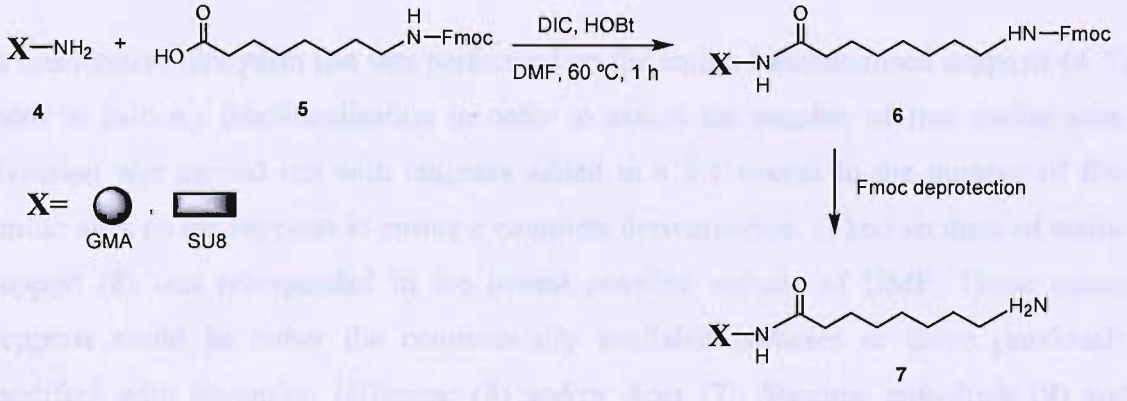
Preparation of amino functionalised SU8 microparticles



1 2 3

Epoxy SU8 microparticles (**1**) (~ 10 mg), were resuspended in acetonitrile (MeCN) (500 μL), and a solution of *bis*-amino Jeffamine (**2**) (500 μL, 0.587 mmol) in MeCN (500 μL) was added. The reaction mixture was agitated at 50 °C for 16 hours. The particles were washed four times with 1 mL of MeCN and four times with 1 mL of methanol (MeOH) to produce amino functionalised SU8 (**3**). The completeness of the reaction was monitored by ninhydrin test.[60]

Amino-octanoic (Aoct) acid derivatisation

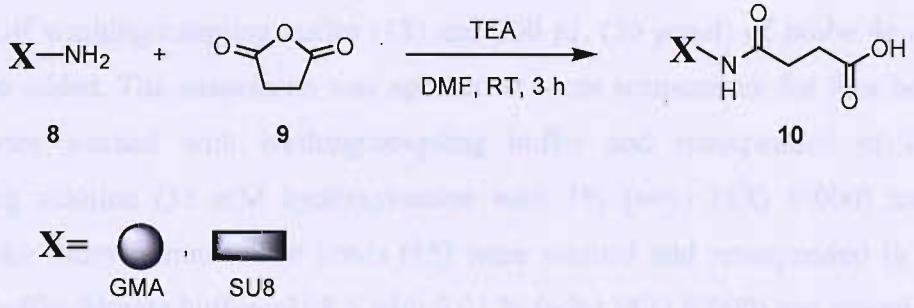


4 5 6 7

A typical derivatisation involved Fmoc – Aoct (**5**), (8.3 mg) dissolved in the lowest possible volume of DMF (typically 20 μL). An aliquot of 20 μL of a DIC solution (32 μL of DIC in 180 μL of DMF) was added and the mixture was agitated for fifteen

minutes. A 10 µL aliquot of a HOBt solution (28 mg in 100 µL of DMF) was added and the mixture was agitated for further ten minutes. The final mixture was poured onto the supports (**4**) and allowed to react for one hour at 60 °C. The supports were washed thoroughly with DMF (**6**) and a qualitative ninhydrin was carried out to assess the level of free amino groups. A positive test meant the repetition of the procedure. If a negative test developed the supports were further capped to ensure that any remaining amino groups would be rendered inaccessible. Capping solutions were made with Tetrahydrofuran (THF) (80%) plus Triethylamine (TEA) (10%) and acetic anhydride (10%) and were mixed with the supports for fifteen minutes twice. Fmoc deprotection (**7**) was achieved with a solution of 20% piperidine in DMF for twenty minutes revealing the amino groups. At this stage a quantitative ninhydrin test was conducted to compare with the Fmoc test results. This procedure could be repeated in order to add consecutive Aoct spacers to the free deprotected amino groups.

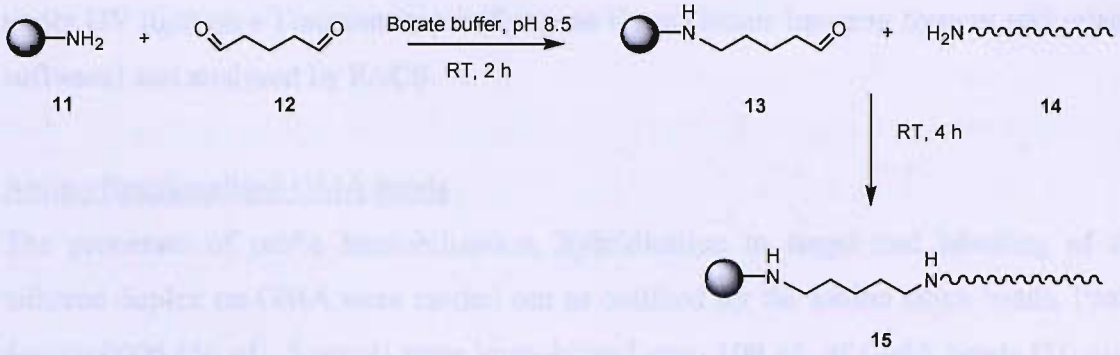
Preparation of carboxy functionalised solid supports



A quantitative ninhydrin test was performed on the amino functionalised supports (**4**, **7**) prior to carboxy functionalisation in order to assess the number of free amino sites. Reaction was carried out with reagents added in a 3:1 excess to the number of free amino sites on the supports to ensure a complete derivatisation. A known mass of amino support (**8**) was resuspended in the lowest possible volume of DMF. These amino supports could be either the commercially available particles or those previously modified with *bis*-amino Jeffamine (**3**) and/or Aoct (**7**). Succinic anhydride (**9**) and TEA, both in a three fold excess, were mixed with the supports for three hours at room temperature with agitation. The supports (**10**) were washed thoroughly with DMF and a

qualitative ninhydrin test was carried out to assess the final results. If a positive test developed at this stage, the entire procedure would be carried out again.

2.5.5 Single base chain extension preliminary experiments



Silica beads (**11**) (600 µL, 30 mg – 220 µmol amino groups per gram of beads) were washed twice in washing/coupling buffer (Borate buffer, pH 8.5) and resuspended in 400 µL of a 10% glutaraldehyde solution (**12**) in the same buffer. The suspension was agitated for two hours at room temperature. It was then washed twice and resuspended in 1 mL of washing/coupling buffer (**13**) and 200 µL (20 µmol) of probe 4goligo0006 (**14**) were added. The suspension was agitated at room temperature for four hours. The beads were washed with washing/coupling buffer and resuspended in 1 mL of quenching solution (35 mM hydroxylamine with 1% (w/v) PEG 6'000) and gently agitated for thirty minutes. The beads (**15**) were washed and resuspended in 1 mL of storage buffer (Borate buffer, pH 8.5 with 0.01 % (w/v) PEG 6'000) and stored at 4 °C.

Hybridisation to target

The beads (**15**) were washed and resuspended in 500 µL of washing/coupling buffer and 100 µL (10 µmol) of target 4goligo0004 were added. The suspension was heated to 94 °C and cooled to room temperature with occasional agitation.[34] The beads were then washed three times with washing/coupling buffer. The absorbance (260 nm) of the supernatants of the coupling and hybridisation reactions was measured (Cary, 1 Bio, UV-Visible Spectrophotometer, Varian).

Single base chain extension

Single base chain extension incubation reactions in a final volume of 50 µL contained the beads with immobilised duplex prepared in the previous step, fluorescein-ddATP (50 µM), $MgCl_2$ (2.5 mM) and *Taq* polymerase (0.5 U) in manufacturer's buffer (1x). The suspension was incubated at 60 °C for two hours with occasional agitation. The beads were washed twice and resuspended in washing/storage buffer and examined under UV light on a Transluminator (Syngene Gene Genius Imaging System and related software) and analysed by FACS.

Amino functionalised GMA beads

The processes of probe immobilisation, hybridisation to target and labelling of the tethered duplex on GMA were carried out as outlined for the amino silica beads. Probe 4goligo0006 (50 µL, 5 µmol) were immobilised onto 100 µL of GMA beads (10 mg - 10 µmol amino functional groups per gram of beads). Hybridisation to target 4goligo0004 was carried out as described and chain extension reaction was incubated for three hours.

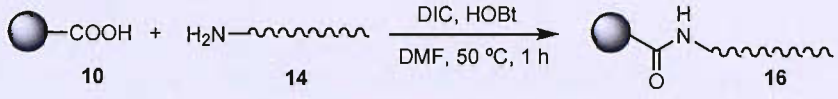
2.5.6 Influence of temperature and mismatches on single base chain extensions

The preparation of amino GMA beads (30 µL, 3 mg) with immobilised probe 4goligo0006 (200 µL - 20 µmol) was as described. Three sets of three microtubes were prepared containing each 100 µL of the suspension of beads. The first and second microtubes of each set contained 25 µL (2.5 µmol) of targets 4goligo0004 and 4goligo0011, respectively. The third tube contained no target. All suspensions underwent hybridisation incubation as described. The reagents for single base chain extension were added to the tubes and each set of tubes was incubated for three hours at 60, 70 and 80 °C, respectively. Suspensions were further incubated overnight at room temperature. The beads were washed twice and resuspended in storage buffer. All suspensions were exposed to UV light and analysed by FACS. The experiment was repeated as described using *Pfu Turbo* DNA polymerase (1.75 U).

2.5.7 Checking for non-specific binding

Three sets of four microtubes containing each approximately 2 mg of amino functionalised GMA, amino functionalised silica, carboxy functionalised GMA and carboxy functionalised silica beads were prepared. To the first set of microtubes, all the beads were incubated with *Taq*, PCR buffer, $MgCl_2$ and fluorescein – ddATP as described in the single base chain extension section. The second set of microtubes contained all reagents except the enzyme. Both sets of tubes were incubated at 60 °C for three hours and overnight at room temperature. The beads were washed three times and resuspended in PBS buffer (10 mM K_2HPO_4, 1.8 mM NaH_2PO_4, 0.15 M NaCl, pH 7.4) and were exposed to UV light.

2.5.8 Assessing efficiency of probe attachment and hybridisation



Carboxy GMA beads (6 mg - 10 μmol amino functional groups per gram of beads) (**10**) were mixed with 50 μL of DMF peptide grade. A 10 μL aliquot of a solution 10% DIC in DMF peptide grade was added and the mixture was agitated at room temperature for 15 minutes. A solution made of 1 mg of HOBt in 20 μL of DMF peptide grade was added and the mixture was agitated for further 10 minutes. The beads were washed twice with DMF peptide grade and four times with acetonitrile. The last wash was kept for analysis (sample A). Probe 4goligo0014 (**14**) (225 μL, 90 nmol in acetonitrile) was added and the mixture was heated at 50 °C for an hour with occasional agitation. The supernatant was taken aside for analysis (sample B). The beads were washed six times and resuspended in 1 mL of TE buffer (10 mM tris-HCl, 1 mM EDTA) (**16**). A 20 μL suspension aliquot was taken for analysis (sample C). To 100 μL of the previous suspension, 5' Cy5 target 4goligo0016 (90 μL, 9nmol) prepared in TE buffer was added and the mixture was incubated to promote hybridisation as described. The supernatant was removed (sample D) and the beads were washed with TE buffer. An aliquot of hybridised beads was taken for analysis (sample E). For the competition experiment non-labelled target 4goligo0015 (90 μL, 9nmol) was added and the re-hybridisation

incubation was carried out. The suspension of beads (sample F) was analysed by FACS. For the labelling experiment, 300 μL of probe immobilised suspension (**16**) were hybridised overnight to target 4goligo0015 (270 μL, 27 nmol). The extension incubation contained in a final volume of 120 μL, the beads with the immobilised duplex, Thermo Sequenase™ DNA polymerase (96 U), Thermo Sequenase dilution buffer (25μx) and fluorescein – 12 – ddATP (25 μM) in manufacturer's buffer (83μx). The mixture was incubated at 60 °C overnight after which the beads were washed and analysed by FACS. Samples A to F were analysed by FACS, HPLC and their absorbance at 260nm (NanoDrop – ND 1000, spectrophotometer) was measured. HPLC analysis was carried out on a HP Agilent 1100 (Column Phenomenex - Reverse Phase, Jupiter, 5 μm, C18, 300Å, 50x4.6 mm). Injection volume 100 μL - UV detection at 254 nm - The mobile phase was a mixture of water/TFA (0.1 %) and acetonitrile/ TFA (0.1 %), starting with 70 % aqueous and 30 % organic for three minutes at 1 ml/min, followed by a three minutes linear gradient to 100 % organic. This was maintained for one minute, followed by a return to 100 % aqueous and 0 % organic. Each injection was complete in 33 minutes.

2.5.9 Single base chain extension in solution

For the labelling carried out in solution, probe 4goligo0014 (100 μL, 10 nmol) was hybridised to target 4goligo0015 (100 μL, 10 nmol) as described. Single base chain extension incubation reactions in a final volume of 64 μL contained 50 μL of the hybridised solution, $MgCl_2$ (*Taq*) (0.8x), *Taq* / Thermo Sequenase™ (5 U / 48 U), fluorescein – 12 – ddATP (47 μM), in manufacturer's buffer (0.8x / 78 μx). Samples were incubated at 60 °C for two hours. Non-denaturing PAGE analysis of these two samples was carried out (Table 2.15).

Reagent	Volume
TBE buffer	2.25 mL
Acrylamide	4.5 mL
10 % ammonium persulphate	70 μL
TEMED	5 μL

Table 2.15 Resolving gel mixture

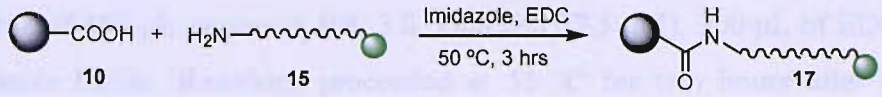
Approximately 2 μL of each DNA sample were mixed with equal quantity of Orange G (2 % Orange G in 10 % glycerol in formamide) and pipetted into the gel lanes. Gels were developed in SYBR® Green RNA for approximately 40 minutes.

2.5.10 Restriction endonuclease digestion

An aliquot of 100 μL of probe immobilised beads (**16**) was hybridised to 5' Cy5 target 4goligo0016 (90 μL, 9nmol) as previously described. The beads were washed and resuspended in 220 μL of TE buffer and an aliquot of 20 μL was analysed by FACS. The remainder 200 μL were washed and resuspended in 100 μL of deionised water. To this suspension 1.5 μL (30 units) of *Xho I*, 10 μL of 10x NEB2 buffer were added and the mixture was incubated at 37 °C for two hours. The beads were washed and analysed by FACS.

2.5.11 Spacers as a mean to improve hybridisation efficiency and enzyme accessibility

Three sets of eight microtubes each were prepared. All microtubes contained 2 mg of carboxy GMA beads (**10**). The beads on the first, second and third sets were modified with 1Aoct, 2Aoct and 3 Aoct (**7**) respectively, prior to their carboxy funtionalisation. The loading level of the beads, as determined by quantitative ninhydrin and Fmoc tests was 6.7, 4.8 and 4.2 μmol per gram of beads, respectively. Each set of microtubes contained in a final volume of 267, 248 and 242 μL, respectively, increasing quantity of probe 4goligo0017 (**15**) (0, 0.3, 1.3, 2.5, 5.0, 7.5, 12.5, 25 μM; 0, 0.2, 0.9, 1.9, 3.9, 5.8, 9.7, 19.3 μM; 0, 0.2, 0.9, 1.7, 3.5, 5.2, 8.7, 17.4 μM) (molar ratio of labelled probe / amount of free amino groups), 200 μL of EDC (100mM of EDC in 100mM imidazole pH 7.0) and imidazole (100mM imidazole pH 7.0). Suspensions were incubated at 50 °C for five hours and with constant agitation and cooled to room temperature. The beads (**17**) were washed, resuspended in 400 μL of deionised water and analysed by FACS.



The attachment of 3' fluorescein probe 4goligo0017 to the beads was carried out in aqueous media in the presence of EDC as a coupling agent. The carbodiimide chemistry of attachment was changed from non-aqueous to aqueous media as the labelled probe was only soluble in the latter.

FRET as confirmation of hybridisation

Aliquots of 200 µL of suspensions of beads with probe loadings of 1, 5 and 10 % were hybridised to 5' Cy5 target, 4goligo0016 (100 µL - 10 nmol) and to non-labelled target, 4goligo0015 (100 µL - 10 nmol), respectively. The beads were washed and resuspended in PBS buffer and analysed by FACS.

Influence of spacers on enzyme accessibility

Aliquots of 200 µL of suspensions of beads with probe loadings of 1, 5, 10 and 20 % were washed and resuspended in 80 µL of deionised water. Digestion incubations for probe loadings of 1, 5 and 10 % contained in a final volume of 103 µL, the beads, Mung Bean exonuclease (30 U) in manufacturer's buffer (2x). Digestion incubation for probe loading of 20 % contained in a final volume of 100 µL, the beads, Mung Bean exonuclease (50 U) in manufacturer's buffer (1.5x). All microtubes were incubated at 30 °C for thirty minutes. The beads were washed and resuspended in PBS buffer and analysed by FACS.

2.5.12 Influence of increasing probe loading on the fluorescence of SU8 microparticles and hybridisation to tethered probes

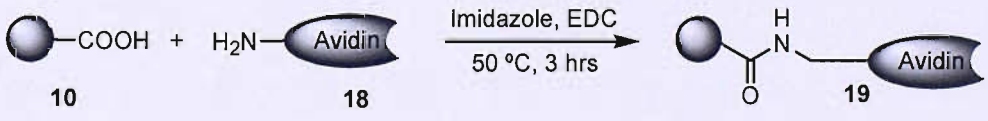
Carboxy functionalised SU8 microparticles (**10**) were obtained after modification with *bis*-amino Jeffamine (**3**) and Aoct (**7**). Fmoc and ninhydrin quantitative tests yielded 18 µmol free amino sites per gram of microparticles. Three sets of five microtubes containing each 1 mg of microparticles were loaded with increasing quantities (0, 2, 10, 50 and 100 %) of probes 4goligo0017, 4goligo0030 and 4goligo0014 (molar ratio of labelled probe / amount of free amino groups). Immobilisation reactions contained in a

final volume of 480 µL, probe (0, 0.8, 3.8, 18.8 and 37.5 µM), 300 µL of EDC solution and imidazole buffer. Reactions proceeded at 50 °C for two hours after which the microparticles were washed and resuspended in 1 mL of an aqueous solution of 0.02 % Tween in order to overcome the inherent stickiness of SU8 and analysed by FACS. For the hybridisation, aliquots of 50 µL of the suspensions prepared were incubated with targets 4goligo0016 (30 µL, 3 nmol) and 4gmismatch (30 µL, 3 nmol), respectively. 20 µL of 20x SSC buffer was added to the suspensions and the beads were heated to 45 °C and cooled to room temperature with occasional agitation. The particles were washed and resuspended in Tris buffer (0.5 M tris-HCl, 1 M NaCl, pH 8.1, 0.02 % Tween) and analysed by FACS.

2.5.13 Investigating the EDC mediated coupling chemistry

Three sets of six microtubes containing each 1 mg (6.7 µmol amino functional groups per gram of beads) of carboxy GMA beads (**10**) previously modified with 1Aoct (**7**) were prepared. The first and second sets of microtubes contained in a final volume of 267 µL, increasing quantities (0, 0.5, 2.5, 12.5 and 25 µM) of probes 4goligo0017 and 4goligo0019, respectively, 200 µL of EDC solution and imidazole buffer. The third set was prepared as the first set but contained no EDC (volume was completed using imidazole). All suspensions were agitated at 50 °C for three hours. The beads were washed and resuspended in Tris buffer (50 mM tris-HCl, 50 mM NaCl, pH 8.1) and analysed by FACS.

2.5.14 Preparation of avidin modified GMA beads

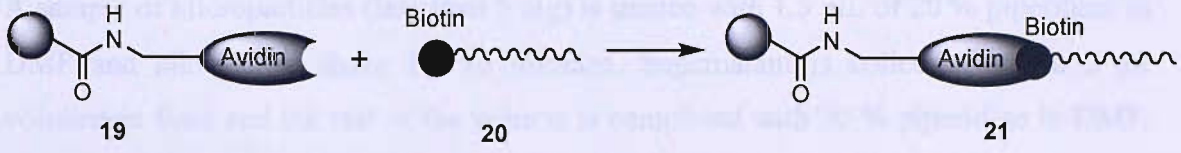


10 18 19

Carboxy GMA beads (**10**) (3,3 mg, 6.7 µmol amino functional groups per gram of beads), without any spacer modification were resuspended in 600 µL of coupling buffer (80 mM EDC in 100 mM imidazole buffer, pH 7.0). Avidin (**18**) (500 µL, 0.5 mg) was

added and the final mixture was agitated at 50 °C for three hours. The beads (**19**) were then washed and resuspended in 1 mL of hybridisation buffer, 5xSSPE (0.75 M NaCl, 50 mM NaH_2PO_4, 5 mM EDTA, pH 7.4), 0.1 % SDS and stored at 4 °C.

2.5.15 Attachment of 5' biotin oligonucleotides to avidin modified GMA beads



19 20 21

Biotinylated probes (0.11 nmol), ATD0747 and ATD0748 (**20**) were added to avidin coated GMA beads (50 µL, 0.166 mg). The mixtures were agitated for thirty minutes, washed and resuspended in 100 µL of hybridisation buffer (**21**).

2.5.16 DNA hybridisation on avidin coated GMA beads and post-hybridisation stringency

GMA beads with immobilised DNA probes (**21**) aliquots, (50 µL, 0.083 mg) were mixed with a tenfold excess of DNA target, (ATD0749, ATD0750 or 4goligo0016) in 200 µL of hybridisation buffer and reactions were incubated at 94 °C for two minutes and overnight at 45 °C with constant agitation. After hybridisation, the beads were washed with buffer solutions of increased stringency; 5xSSPE, 0.1%, SDS, 2xSSC (0.3 M NaCl, 30 mM Na_3citrate, pH 7.0), 0.1%, SDS and 1xSSC (0.15 M NaCl, 15 mM Na_3citrate, pH 7.0), 0.1% SDS, and analysed by FACS.

2.5.17 UV and Colorimetric tests

Fmoc Test

Loading level (mmol/g) = [Abs (302) V (ml) 1000] / [ε (302) w (mg)]
Extinction coefficient, ε (302) = 7800 M^{-1} cm^{-1}

A sample of microparticles (less than 5 mg) is treated with 1.5 mL of 20 % piperidine in DMF and allowed to shake for 15 minutes. Supernatant is collected into a 5 ml volumetric flask and the rest of the volume is completed with 20 % piperidine in DMF. Absorbance is measured at 302 nm against a blank of 20 % piperidine in DMF.

Ninhydrin Test

Loading level (mmol/g) = [Abs (570) V (mL) 1000] / [ε (570) w (mg)]
Extinction coefficient, ε (570) = 15000 M^{-1} cm^{-1}

Quantitative Ninhydrin test – A glass test tube containing a sample of microparticles (less than 5 mg) is treated with 6 droplets of Kaiser solution A (3.86 M phenol in absolute EtOH, 0.36 mM KCN, 11 mM pyridine) and 3 droplets of Kaiser solution B (280 mM ninhydrin in absolute EtOH) and heated at 110 °C for 10 minutes. The sample is transferred into a 5 mL volumetric flask after cooling down and the glass tube washed with 60% EtOH adding the washes to volumetric flask. The final volume is completed with 60% EtOH and the microparticles allowed settling down. Absorbance is measured at 570 nm against a blank of 60% EtOH.

Qualitative Ninhydrin test - A glass test tube containing a sample of microparticles (less than 5 mg) is treated with 6 droplets of Kaiser solution A and 3 droplets of Kaiser solution B and heated at 110 °C for 10 minutes, if a blue colour develops then the sample contains amine groups.

2.6 References

1. Bradley M. et al., Basic Technology grant proposal - 4 billion bases a day - practical individual genome sequencing. **2002**.
2. Liu, J.; Cai, B.; Zhu, J.; Ding, G.; Zhao, X.; Yang, C.; Chen, D., Process research of high aspect ratio microstructure using SU-8 resist. *Microsystem Technologies* **2004**, 10, 265-8.
3. Cavalli, G.; Banu, S.; Ranasinghe, R.T.; Broder, G.; Martins, H.F.P.; Neylon, C.; Morgan, H.; Bradley, M.; Roach, P., Multistep synthesis on SU-8: Combining microfabrication and solid-phase chemistry on a single material. **submitted**.
4. Marie, R.; Schmid, S.; Johansson, A.; Ejsing, L.; Nordstrom, M.; Hafliger, D.; Christensen, C.B.; Boisen, A.; Dufva, M., Immobilisation of DNA to polymerised SU-8 photoresist. *Biosens. Bioelectron.* **2006**, 21, 1327-32.
5. Holmes, D.; Whitton, A.; Cavalli, G.; Banu, S.; Birtwell, S.; Galitonov, G.; Martins, H.; Neylon, C.; Zheludev, N.; Morgan, H., Optical and impendance detection for analysis of DNA hybridisation on encoded micro-particles. Micro Total Analysis Systems **2006**, Proceedings. Tokyo, Japan: Kluwer Academic Publ. 1306-8.
6. Walsh, M.K.; Wang, X.W.; Weimer, B.C., Optimizing the immobilization of single-stranded DNA onto glass beads. *J. Biochem. Biophys. Methods* **2001**, 47, 221-31.
7. Brenner, S.; Johnson, M.; Bridgham, J.; Golda, G.; Lloyd, D.H.; Johnson, D.; Luo, S.J.; McCurdy, S.; Foy, M.; Ewan, M.; Roth, R.; George, D.; Eletr, S.; Albrecht, G.; Vermaas, E.; Williams, S.R.; Moon, K.; Burcham, T.; Pallas, M.; DuBridge, R.B.; Kirchner, J.; Fearon, K.; Mao, J.; Corcoran, K., Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. *Nat. Biotechnol.* **2000**, 18, 630-4.
8. Brenner, S.; Williams, S.R.; Vermaas, E.H.; Storck, T.; Moon, K.; McCollum, C.; Mao, J.I.; Luo, S.J.; Kirchner, J.J.; Eletr, S.; DuBridge, R.B.; Burcham, T.; Albrecht, G., In vitro cloning of complex mixtures of DNA on microbeads: Physical separation of differentially expressed cDNAs. *Proc. Natl. Acad. Sci. USA* **2000**, 97, 1665-70.
9. Kohara, Y.; Noda, H.; Okano, K.; Kambara, H., DNA probes on beads arrayed in a capillary, 'Bead-array', exhibited high hybridization performance. *Nucleic Acids Res.* **2002**, 30, e87.
10. Steinberg, G.; Stromsborg, K.; Thomas, L.; Barker, D.; Zhao, C.F., Strategies for covalent attachment of DNA to beads. *Biopolymers* **2004**, 73, 597-605.
11. Watterson, J.; Piunno, P.A.E.; Krull, U.J., Practical physical aspects of interfacial nucleic acid oligomer hybridisation for biosensor design. *Anal. Chim. Acta* **2002**, 469, 115-27.
12. Charreyre, M.T.; Tcherkasskaya, O.; Winnik, M.A.; Hiver, A.; Delair, T.; Cros, P.; Pichot, C.; Mandrand, B., Fluorescence energy transfer study of the conformation of oligonucleotides covalently bound to polystyrene latex particles. *Langmuir* **1997**, 13, 3103-10.
13. Herne, T.M.; Tarlov, M.J., Characterization of DNA probes immobilized on gold surfaces. *J. Am. Chem. Soc.* **1997**, 119, 8916-20.
14. Steel, A.B.; Herne, T.M.; Tarlov, M.J., Electrochemical quantitation of DNA immobilized on gold. *Anal. Chem.* **1998**, 70, 4670-77.
15. Steel, A.B.; Herne, T.M.; Tarlov, M.J., Electrostatic interactions of redox cations with surface- immobilized and solution DNA. *Bioconjug. Chem.* **1999**, 10, 419-23.

16. Adessi, C.; Matton, G.; Ayala, G.; Turcatti, G.; Mermod, J.J.; Mayer, P.; Kawashima, E., Solid phase DNA amplification: characterisation of primer attachment and amplification mechanisms. *Nucleic Acids Res.* **2000**, 28, e87.
17. Beattie, K.L.; Beattie, W.G.; Meng, L.; Turner, S.L.; Coralvazquez, R.; Smith, D.D.; McIntyre, P.M.; Dao, D.D., Advances in Genosensor Research. *Clin. Chem.* **1995**, 41, 700-6.
18. Beattie, W.G.; Meng, L.; Turner, S.L.; Varma, R.S.; Dao, D.D.; Beattie, K.L., Hybridization of DNA targets to glass-tethered oligonucleotide probes. *Mol. Biotechnol.* **1995**, 4, 213-25.
19. Ghosh, S.S.; Musso, G.F., Covalent attachment of oligonucleotides to solid supports. *Nucleic Acids Res.* **1987**, 15, 5353–72
20. Lund, V.; Schmid, R.; Rickwood, D.; Hornes, E., Assessment of Methods for Covalent Binding of Nucleic-Acids to Magnetic Beads, Dynabeads, and the Characteristics of the Bound Nucleic-Acids in Hybridization Reactions. *Nucleic Acids Res.* **1988**, 16, 10861-80.
21. Ali, M.F.; Kirby, R.; Goodey, A.P.; Rodriguez, M.D.; Ellington, A.D.; Neikirk, D.P.; McDevitt, J.T., DNA hybridization and discrimination of single-nucleotide mismatches using chip-based microbead arrays. *Anal. Chem.* **2003**, 75, 4732-39.
22. Cai, H.; White, P.S.; Torney, D.; Deshpande, A.; Wang, Z.L.; Keller, R.A.; Marrone, B.; Nolan, J.P., Flow cytometry-based minisequencing: A new platform for high- throughput single-nucleotide polymorphism scoring. *Genomics* **2000**, 66, 135-43.
23. Pirrung, M.C.; Worden, J.D.; Labriola, J.P.; Montague-Smith, M.P.; Weislo, L.J., Solid-phase, single nucleotide primer extension of DNA/RNA hybrids by reverse transcriptases. *Bioorg. Med. Chem. Lett.* **2001**, 11, 2437-40.
24. Chen, J.W.; Iannone, M.A.; Li, M.S.; Taylor, J.D.; Rivers, P.; Nelsen, A.J.; Slentz-Kesler, K.A.; Roses, A.; Weiner, M.P., A microsphere-based assay for multiplexed single nucleotide polymorphism analysis using single base chain extension. *Genome Res.* **2000**, 10, 549-57.
25. Larson, B.J.; Helgren, J.M.; Manolache, S.O.; Lau, A.Y.; Lagally, M.G.; Denes, F.S., Cold-plasma modification of oxide surfaces for covalent biomolecule attachment. *Biosens. Bioelectron.* **2005**, 21, 796-801.
26. Chiu, S.K.; Hsu, M.; Ku, W.C.; Tu, C.Y.; Tseng, Y.T.; Lau, W.K.; Yan, R.Y.; Ma, J.T.; Tzeng, C.M., Synergistic effects of epoxy- and amine-silanes on microarray DNA immobilization and hybridization. *Biochem. J.* **2003**, 374, 625-32.
27. Goto, S.; Takahashi, A.; Kamisango, K.; Matsubara, K., Single-nucleotide polymorphism analysis by hybridization protection assay on solid support. *Anal. Biochem.* **2002**, 307, 25-32.
28. Peterson, A.W.; Heaton, R.J.; Georgiadis, R., Kinetic control of hybridization in surface immobilized DNA monolayer films. *J. Am. Chem. Soc.* **2000**, 122, 7837-38.
29. Schmitz, C.; Reetz, M.T., Solid-phase enzymatic synthesis of oligonucleotides. *Org. Lett.* **1999**, 1, 1729-31.
30. Shchepinov, M.S.; CaseGreen, S.C.; Southern, E.M., Steric factors influencing hybridisation of nucleic acids to oligonucleotide arrays. *Nucleic Acids Res.* **1997**, 25, 1155-61.
31. Peterson, A.W.; Wolf, L.K.; Georgiadis, R.M., Hybridization of mismatched or partially matched DNA at surfaces. *J. Am. Chem. Soc.* **2002**, 124, 14601-7.
32. Laboratories, B., TechNote 205 Rev. 003, Active: 30/MAR/02 - Covalent Coupling, Published by Bangs Laboratories, Inc.
33. Laboratories, B., TechNote 302 Rev. 002, Active: 5/FEB/02 - Molecular Biology, Published by Bangs Laboratories, Inc.

34. Wilkins Stevens, P.; Henry, M.R.; Kelso, D.M., DNA hybridization on microparticles: determining capture-probe density and equilibrium dissociation constants. *Nucleic Acids Res.* **1999**, 27, 1719-27.
35. Clark, J. M., Novel non-templated nucleotide addition reactions catalyzed by procaryotic and eucaryotic DNA polymerases. *Nucleic Acids Res.* **1988**, 16, 9677-86.
36. http://www.eppendorf.com/en/supp/applications/applic22en.pdf#search =%22 pfu%20turbo%22.
37. Odonnell, M.J.; Tang, K.; Koster, H.; Smith, C.L.; Cantor, C.R., High density, covalent attachment of DNA to silicon wafers for analysis by MALDI-TOF mass spectrometry. *Anal. Chem.* **1997**, 69, 2438-43.
38. Evans, S.J.; Fogg, M.J.; Mamone, A.; Davis, M.; Pearl, L.H.; Connolly, B.A., Improving dideoxynucleotide-triphosphate utilisation by the hyper-thermophilic DNA polymerase from the archaeon Pyrococcus furiosus. *Nucleic Acids Res.* **2000**, 28, 1059-66.
39. Arezi, B.; Hansen, C.J.; Hogrefe, H.H., Efficient and high fidelity incorporation of dye-terminators by a novel archaeal DNA polymerase mutant. *J. Mol. Biol.* **2002**, 322, 719-29.
40. Tabor, S.; Richardson, C.C., A single residue in DNA polymerases of the Escherichia coli DNA polymerase I family is critical for distinguishing between deoxy- and dideoxyribonucleotides. *Proc. Natl. Acad. Sci. USA* **1995**, 92, 6339-43.
41. Brandis, J.W., Dye structure affects Taq DNA polymerase terminator selectivity. *Nucleic Acids Res.* **1999**, 27, 1912-8.
42. Li, Y.; Mitaxov, V.; Waksman, G., Structure-based design of Taq DNA polymerases with improved properties of dideoxynucleotide incorporation. *Proc. Natl. Acad. Sci. USA* **1999**, 96, 9491-6.
43. Watterson, J.H.; Piunno, P.A.E.; Krull, U.J., Towards the optimization of an optical DNA sensor: control of selectivity coefficients and relative surface affinities. *Anal. Chim. Acta* **2002**, 457, 29-38.
44. Selvin, P.R., The renaissance of fluorescence resonance energy transfer. *Nat. Struct. Biol.* **2000**, 7, 730-4.
45. Okamura, Y.; Watanabe, Y., Detecting RNA/DNA hybridization using double-labeled donor probes with enhanced fluorescence resonance energy transfer signals. *Methods Mol. Biol.* **2006**, 335, 43-56.
46. Wahlroos, R.; Toivonen, J.; Tirri, M.; Hanninen, P., Two-Photon Excited Fluorescence Energy Transfer: A Study Based on Oligonucleotide Rulers. *J. Fluoresc.* **2006**, 16, 379-86.
47. Howell, W.M., Detection of DNA hybridization using induced fluorescence resonance energy transfer. *Methods Mol. Biol.* **2006**, 335, 33-41.
48. http://probes.invitrogen.com/resources/spectraviewer/.
49. Crockett, A.O.; Wittwer, C.T., Fluorescein-labeled oligonucleotides for real-time pcr: using the inherent quenching of deoxyguanosine nucleotides. *Anal. Biochem.* **2001**, 290, 89-97.
50. Broder, G. R., *M.Phil/Ph.D. - Transfer Report* - University of Southampton. **2006**.
51. Bunemann, H., Immobilization of denatured DNA to macroporous supports: II. Steric and kinetic parameters of heterogeneous hybridization reactions *Nucleic Acids Res.* **1982**, 10, 7181-96.
52. Bunemann, H.; Westhoff, P.; Herrmann, R.G., Immobilization of denatured DNA to macroporous supports: I. Efficiency of different coupling procedures. *Nucleic Acids Res.* **1982**, 10, 7163-80.

53. Green, N.M., Avidin. 3. The nature of the biotin-binding site. *Biochem. J.* **1963**, 89, 599-609.
54. Green, N.M., Avidin. 4. Stability at extremes of pH and dissociation into sub-units by guanidine hydrochloride. *Biochem. J.* **1963**, 89, 609-20.
55. http://en.wikipedia.org/wiki/Avidin.
56. Gingeras, T.R.; Kwoh, D.Y.; Davis, G.R., Hybridisation properties of immobilized nucleic acids. *Nucleic Acids Res.* **1987**, 15, 5373-90.
57. McQuain, M.K.; Seale, K.; Peek, J.; Fisher, T.S.; Levy, S.; Stremler, M.A.; Haselton, F.R., Chaotic mixer improves microarray hybridization. *Anal. Biochem.* **2004,** 15, 215-26.
58. Sambrook, J.; Fritsch, E.F.; Maniatis, T., Molecular Cloning - A Laboratory Manual. **1989**, Cold Spring Harbor Press, Cold Spring Harbor, New York (2.60-2.80), (2nd edn.).
59. Schaupp, C.J.; Jiang, G.; Myers, T.G.; Wilson, M.A., Active mixing during hybridization improves the accuracy and reproducibility of microarray results. *Biotechniques* **2005**, 38, 117-9.
60. Kaiser, E.; Colescott, R.L.; Bossinger, C.D.; Cook, P. I., Color test for detection of free terminal amino groups in the solid-phase synthesis of peptides. *Anal. Biochem.* **1970**, 34, 595-8.

3. Investigating the selectivity of hybridisation

3. 1 Introduction

Microarrays are based upon hybridisation to form the Watson–Crick duplex as the result of a mixed phase reaction between complementary nucleic acid strands.[1, 2] The structure of the resulting duplex is determined, in part by base pairing and base stacking, in combination with the constraints forced on the phosphodiester backbone conformation and sugar moieties.[2] In addition, the duplex can be affected by the presence and influence of the solid support onto which the DNA probes are attached[3] and by the probe immobilisation method[4, 5] as seen in Chapter 2. Hybridisation kinetics and thermodynamics for duplex formation and dissociation are dependent on oligonucleotide probe length,[6] probe density,[4, 5, 7, 8] base composition (G + C %),[9] number, identity and location of mismatches, temperature,[10-12] ionic strength and the concentration of destabilising agents such as formamide and urea.[13, 14]

3.1.1 The effect of probe length on hybridisation

Short oligonucleotide probes (typically ≤ 25-mers) have the advantage of being accurately synthesised in higher yields when compared to longer probes.[15] Short probes are also cost effective as they require a small number of synthetic steps to be prepared increasing the speed of the probe library preparation.

However, the use of short probes is still hindered by a lack of knowledge of the hybridisation properties of short duplexes in mixed-phase assays. [9, 16] Also, the unavailability of systematic studies on mixed-phase short sequence hybridisation and proper models for interpretation of resultant data, translates into a poor understanding of the limitations of implementing short probes. As a result, the short probes currently

synthesised on microarrays can be ineffective displaying limited sequence specificity or low sensitivity. [15, 17-20] Low specificity translates into a high cross-hybridisation (hybridisation of the probe to the wrong target or to the wrong section of the target) and generation of false positives.[21] Low sensitivity means low signal/noise ratio which results in low detection of the genes under study.

On the other hand, longer oligonucleotide probes show greater sensitivity but are costly to prepare as they require more synthetic steps and are obtained in lower yields. However, longer probes are also prone to cross-hybridisation.[15, 17, 22] Also, increasing the probe length decreases SNP discrimination.[22]

3.1.2 Effects of base composition and sequence on hybridisation

Base composition has a large effect on duplex yield in the solvents normally used for hybridisation. The effect is due to the lower stability of A:T versus G:C base-pairs. Short oligonucleotides may have extreme biases in G:C content and oligonucleotides of the same length can have correspondingly large differences in melting temperature (T_m). Approximately, adding an A:T base pair increases T_m by 2 °C, compared with 4 °C for a G:C pair.[14, 23]

The thermal stability of duplexes formed between DNA sequences with mismatched bases is usually decreased according to the number, identity and location of those mismatches when compared with the corresponding perfectly paired duplex. The hybridisation efficiency of the mismatched oligonucleotide is expected to be generally lower than that of the perfectly matched duplex and consequently, the dissociation rate of the mismatched duplex is expected to be higher than that of the perfectly matched duplex.[14, 24] The stability of mismatched base pairs differs according to the bases involved and with the neighbouring base-pairs. Mismatched base pairs such as G:G, G:T or G:A can be relatively stable, whereas mismatches such as A:A, T:T, C:T, C:A and C:C are usually very unstable and destabilise the duplex structure significantly.[25-32]

3.1.3 Hybridisation stringency

The hybridisation solution conditions are crucial for the optimal performance of the array platform as one of the most significant problems relating to array assays is the control and quantitation of non-specific binding to the oligonucleotides probes. This non-specific binding refers to the lower affinity mismatched duplexes involving sequences other than the complementary target. Microarrays rely thus on the differences in hybridisation stability of oligonucleotides to perfectly matched and mismatched target sequence variants. A system with substantial levels of non-specific binding will in general have reduced ability to detect low abundance sequence species and to discriminate closely related species.[1]

In order to minimise the hybridisation of the tethered probes to related but non-identical sequences, two stringency approaches can be undertaken. The stringency can be adjusted either during the hybridisation step or in the post-hybridisation washes. During the hybridisation step, stringency can be altered by adjusting the salt concentration and/or by changing the temperature. Alternatively hybridisation can be performed at low stringency (high salt concentration and low temperature) and washing at increasing stringencies. The analysis of the results after each wash enables the monitoring of the effectiveness of the washes in hybridisation discrimination. [20, 24]

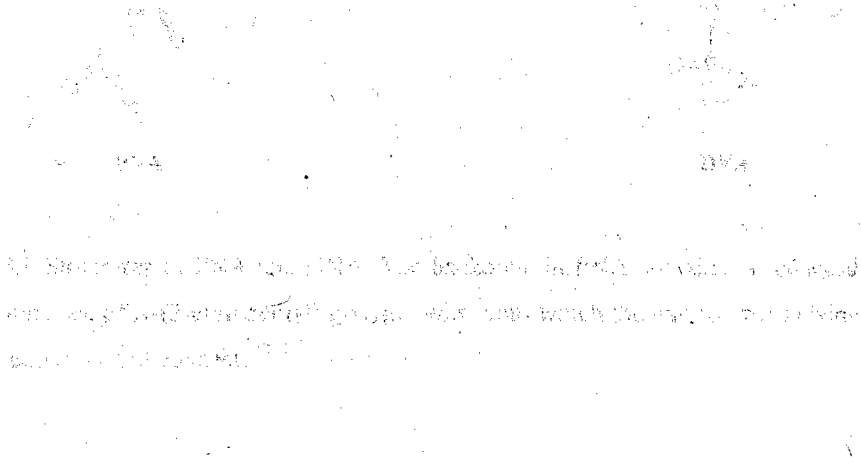
3.2 Aims of this chapter

Chapter 2 showed how the probe immobilisation method can severely affect the hybridisation signal of a suspension array. The ineffective EDC mediated chemistry of probe attachment rendered the oligonucleotide probes less available for hybridisation compromising the efficiency of the platform. The introduction of the avidin-biotin system of probe attachment allied to a revised hybridisation methodology led to an improvement on the hybridisation efficiency and hybridisation discrimination. The combination of a high ionic strength buffer,[14, 33] smaller hybridisation volumes and constant agitation of the beads, allowed both immobilised probes and targets to come together in a more effective way[14, 34-36] overcoming kinetic and diffusion issues and

maximising the rate of annealing. A post-hybridisation washing procedure involving washing the beads with buffers and urea solutions of increased stringency,[14, 24] reducing the non-specific binding and enhancing hybridisation selectivity, was also utilised.

The preliminary results showed hybridisation selectivity and hybridisation discrimination between complementary and non-complementary sequences. These initial results motivated the optimisation of the hybridisation methodology described above and hybridisation discrimination was further investigated for short immobilised probes on GMA beads.

This chapter explores how short probes will affect the specificity of the hybridisation. The effect of probe sequence on the hybridisation discrimination of 16-mers was studied. A SNP detection using 14-mers was also investigated. The potential of multiplexed suspension arrays and their application for interrogating DNA targets using 14-mers, 20-mer and 26-mer probes is also explored. The hybridisation discrimination between complementary and non-complementary targets using DNA and a DNA analogue, PNA, as capture probe was also compared.



3. 3 Results and Discussion

3.3.1 The potential of PNA as a DNA analogue probe

Sequence-specific binding to DNA targets by artificial agents is of major interest for the development of gene-therapeutic, gene-diagnostic, and molecular biology applications.[37, 38] One such agent, peptide nucleic acid (PNA), is a remarkable mimic of nucleic acids (Figure 3.1) in which the phosphodiester backbone has been replaced by an achiral pseudo-peptide chain. PNA binds to complementary DNA and RNA by Watson-Crick type base pairing with remarkably high affinity and selectivity to form duplexes of high stability. These interactions are in general more selective than DNA-DNA or RNA-RNA interactions, due to the lack of the negatively charged backbone which results in a lack of interstrand electrostatic repulsion causing a salt independent hybridisation.[39, 40]

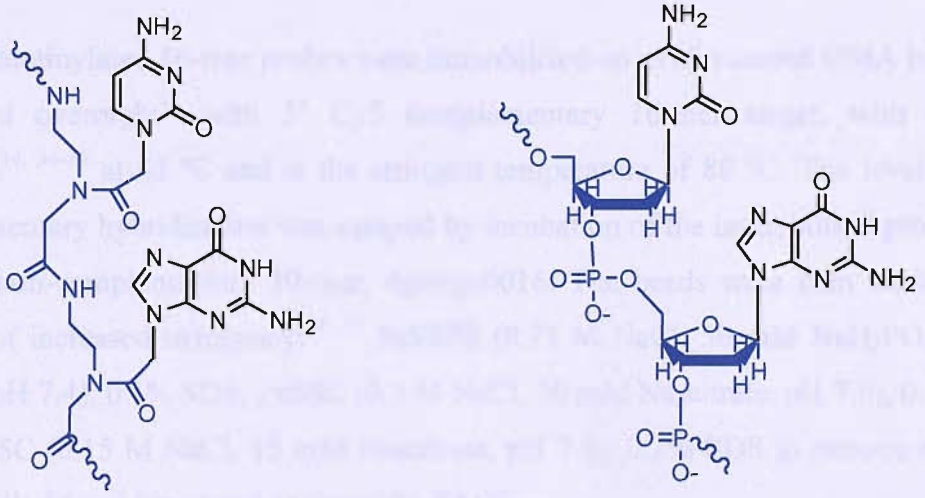


Figure 3.1 Structures of PNA and DNA. The backbone in PNA contains no charged phosphate groups and is composed of *N*-(2-aminoethyl) glycine units, onto which the natural recognising elements of DNA (the nucleobases) will connect. [38, 40]

PNA also possesses the additional advantage of resistance to nucleases and proteolytic degradation.[39, 40] Due to all these properties, PNA has the potential to replace DNA

oligonucleotides in several biotechnology and molecular biology hybridisation applications such as in situ hybridisation arrays, hybridisation sensors and restriction enzyme blocking.[37, 38, 40]

The potential of PNA as a DNA analogue probe for suspension arrays was investigated. The thermal stability of PNA-DNA and DNA-DNA duplexes was compared using probes ATD0748, PNA161 and complementary target ATD0750 (Table 3.1).

Probes	5' modification	Sequence 5' – 3'	Length (nt)
ATD0748	Biotin	CTAGTTACTCTTGTTC	16
PNA161	Biotin	CTAGTTACTCTTGTTC	16
Targets			
ATD0750	Cy5	GAACAAGAGTAACTAG	16
4goligo0016	Cy5	CCATCGACTTGACTTACTCGAGTGCATCTC	30

Table 3.1 DNA and PNA sequences employed for the PNA-DNA and DNA-DNA thermal stability comparison.

Both 5' biotinylated 16-mer probes were immobilised on avidin coated GMA beads and incubated overnight[14] with 5' Cy5 complementary 16-mer target, with constant agitation[14, 34-36] at 45 °C and at the stringent temperature of 80 °C. The level of non-complementary hybridisation was assayed by incubation of the immobilised probes with 5' Cy5 non-complementary 30-mer, 4goligo0016. The beads were then washed with buffers of increased stringency:[14, 33] 5xSSPE (0.75 M NaCl, 50 mM NaH_2PO_4, 5 mM EDTA, pH 7.4), 0.1% SDS, 2xSSC (0.3 M NaCl, 30 mM Na_3citrate, pH 7.0), 0.1% SDS and 1xSSC (0.15 M NaCl, 15 mM Na_3citrate, pH 7.0), 0.1% SDS to remove any non-specifically bound target and analysed by FACS.

Figures 3.2 and 3.3 show the fluorescence obtained for the DNA-DNA and PNA-DNA duplexes and the complementary and non-complementary hybridisation discrimination achieved at 45 °C and at 80 °C respectively. At 45 °C the DNA-DNA duplex seems to exhibit a higher stability than PNA-DNA as shown by the higher fluorescence observed. However, at 80 °C both DNA-DNA and PNA-DNA duplexes appear to have similar stability. The level of non-complementary hybridisation seems to be lower for the PNA-DNA duplex at both temperatures.

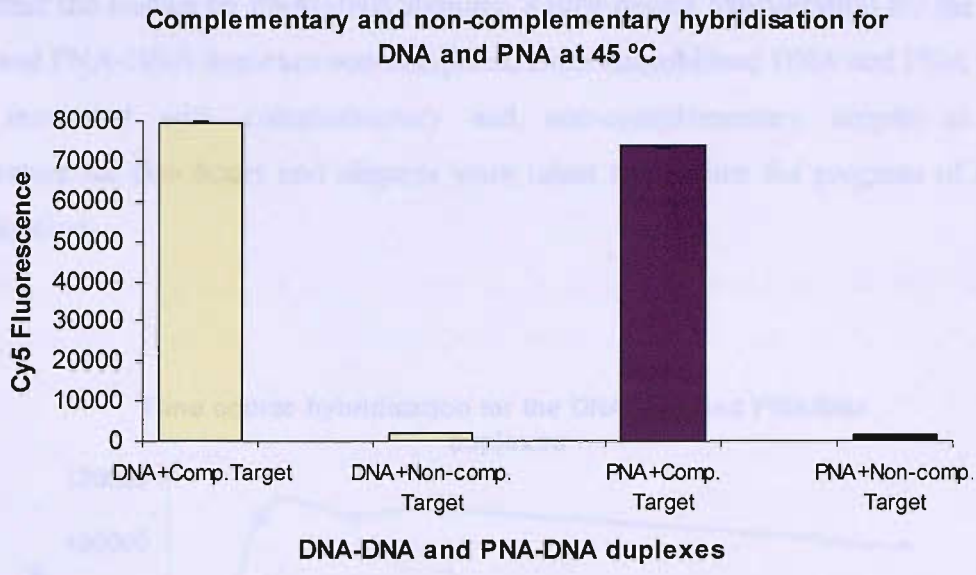


Figure 3.2 Fluorescence (± SE) of complementary and non-complementary hybridisation of DNA-DNA and PNA-DNA duplexes at 45 °C. A control of avidin coated GMA beads incubated with target 4goligo0016 in the same experimental conditions yielded a fluorescence of 633±1.9.

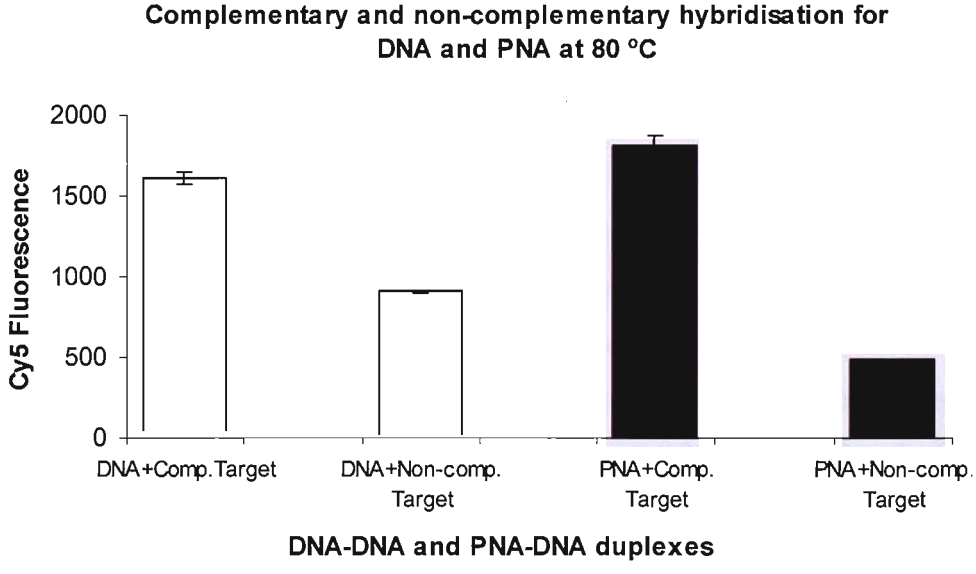


Figure 3.3 Fluorescence (± SE) of complementary and non-complementary hybridisation of DNA-DNA and PNA-DNA duplexes at 80 °C. A control of avidin coated GMA beads incubated with target 4goligo0016 in the same experimental conditions yielded a fluorescence of 418±1.7.

To further the studies on PNA-DNA stability, a time course hybridisation for the DNA-DNA and PNA-DNA duplexes was compared. Both immobilised DNA and PNA probes were incubated with complementary and non-complementary targets at room temperature for two hours and aliquots were taken to monitor the progress of duplex hybridisation.

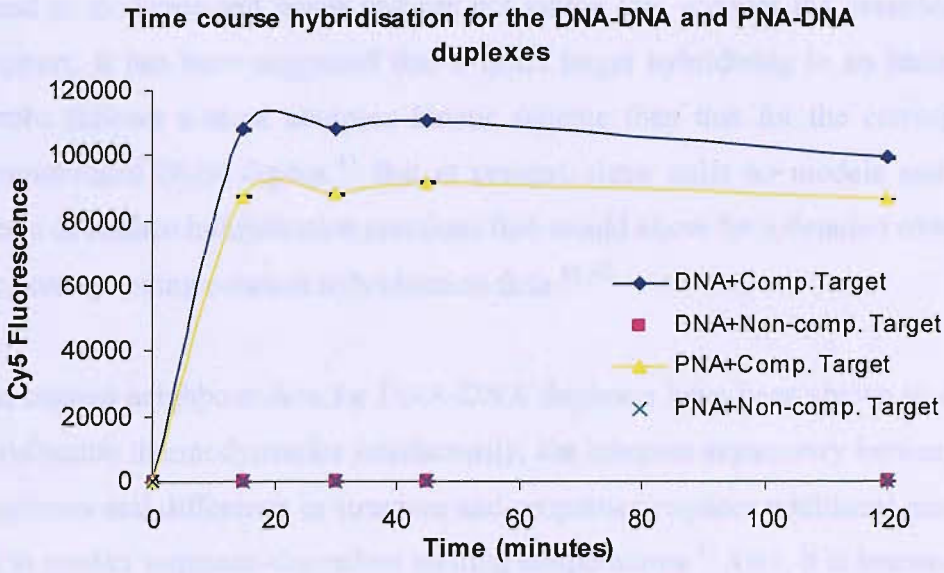


Figure 3.4 Time course hybridisation for DNA-DNA and PNA-DNA duplexes. Fluorescence (± SE) data points are connected to indicate the general trend.

Results shown in Figure 3.4 indicate that hybridisation is readily achieved for both probes fifteen minutes after the start of the experiment. Also, the fluorescence obtained at room temperature is higher than in previous experiments. The levels of non-complementary hybridisation obtained for the time course experiments are lower than those observed for overnight incubations. This suggests that shorter incubation times result in lower non-specific binding. The lower fluorescence intensity observed for the PNA-DNA complementary duplex suggests a lower stability than that of the DNA-DNA complementary pair.

The PNA-DNA duplex is usually more stable than the DNA-DNA duplex.[38, 40] However, for some sequences this statement may not always be true. The studies carried out here show one such situation. For the sequences studied, the DNA-DNA duplex exhibited greater stability than the PNA-DNA duplex.

A possible explanation for the results observed might reside in the fact that the thermodynamic predictions for PNA-DNA duplexes are evaluated for solution like media and at moderate salt levels and are not taking into account the presence of the solid support. It has been suggested that a DNA target hybridising to an immobilised PNA probe follows a more complex kinetic scheme than that for the corresponding DNA-immobilised DNA duplex.[41] But at present, there exist no models and kinetic parameters of surface hybridisation reactions that would allow for a detailed comparison with the corresponding solution hybridisation data.[41, 42]

Whereas nearest neighbour data for DNA-DNA duplexes have been shown to describe the hybridisation thermodynamics satisfactorily, the inherent asymmetry between PNA-DNA duplexes and difference in structure and properties requires additional parameters in order to predict sequence-dependent melting temperatures.[37] Also, it is known that the PNA-DNA duplex stability depends on the purine content[37, 43] of the PNA strand. This property alters the hybridisation efficiency of PNA probes and at a given temperature a fraction of the DNA targets do not obey Watson-Crick binding rules,[43] compromising the viability of the assay.

Another possible explanation might reside in the presence of SDS (sodium dodecyl sulfate) in the hybridisation buffer. SDS is known to work by disrupting non-covalent bonds in the proteins, thereby denaturing them. Also, anions of SDS bind to the main peptide chain and this effectively imparts a negative charge on the protein causing an electrostatic repulsion. Possibly, the backbone of PNA might be affected in a similar manner to the described due to its pseudo-peptide like structure and similar properties explaining the fluorescence results.

3.3.2 The effect of sequence on the hybridisation discrimination of 16-mers

Previous section showed that under the experimental conditions employed, the PNA-DNA duplexes failed to exhibit an improved stability over DNA-DNA duplexes. This lack of success allied to the fact that PNA probes are more expensive to prepare than DNA probes led to the abandonment of hybridisation studies using the DNA analogue.

The experiments carried out in Chapter 2 and in previous section using the avidin-biotin system of probe attachment, demonstrated that tethered 16-mer probes, ATD0747 and ATD0748 had the ability to discriminate against the hybridisation to a non-complementary 30-mer target (4goligo0016). To further the studies on the effect of probe sequence on hybridisation, the ability of each probe to discriminate against the hybridisation to the complementary target of the other probe was investigated. Both 5' biotinylated probes, ATD0747 and ATD0748, were immobilised on avidin coated GMA beads and hybridised to their 5' Cy5 labelled complementary targets (ATD0749 and ATD0750 respectively) and to the complementary target of the other probe. Figure 3.5 summarises the duplexes formed upon hybridisation highlighting the mismatched bases.

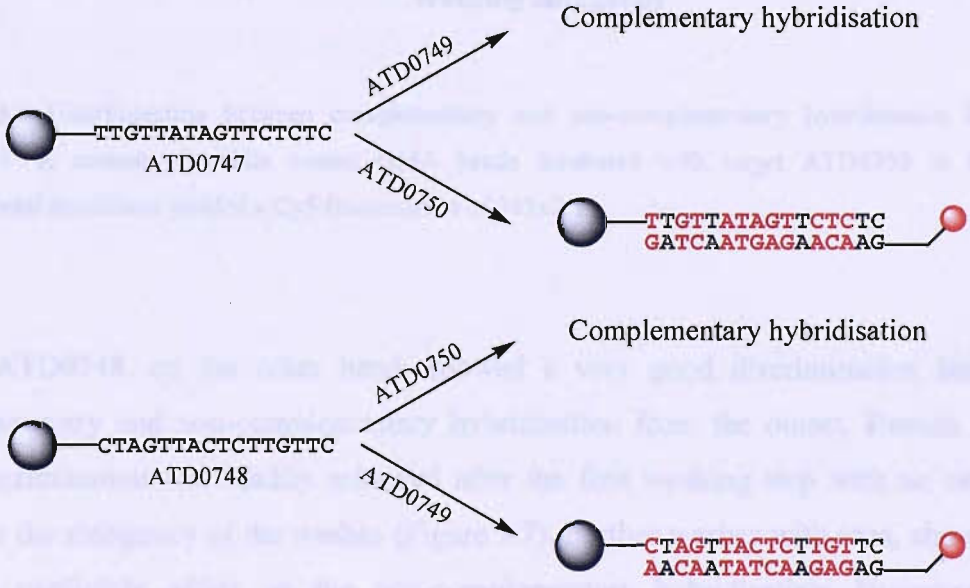


Figure 3.5 DNA duplexes formed upon complementary and non-complementary hybridisations. Mismatches are highlighted in red.

All suspensions were incubated overnight at room temperature. The beads were then washed with buffers of increased stringency (5xSSPE, 0.1% SDS, 2xSSC, 0.1% SDS, 1xSSC, 0.1% SDS) and analysed by FACS at each step.[24] However, probe ATD0747 showed a poor discrimination between complementary and non-complementary hybridisation throughout the washing steps. In order to improve on this poor discrimination, the beads were further washed with a series of urea solutions of increased stringency[14] (50 mM, 1 M and 2 M) but no improvement was observed (Figure 3.6).

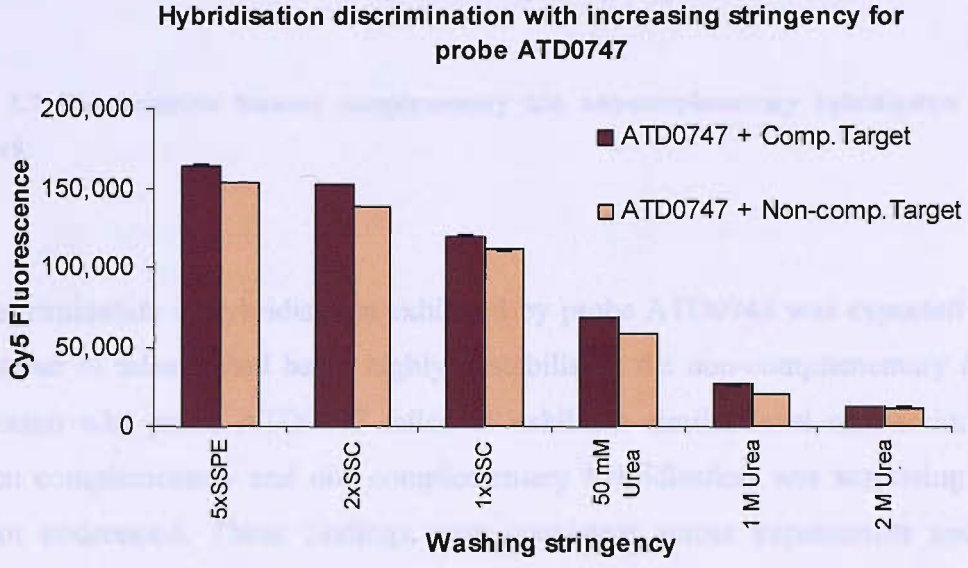


Figure 3.6 Discrimination between complementary and non-complementary hybridisation for probe ATD0747. A control of avidin coated GMA beads incubated with target ATD0750 in the same experimental conditions yielded a Cy5 fluorescence of 347±2.5.

Probe ATD0748, on the other hand, showed a very good discrimination between complementary and non-complementary hybridisation from the outset. Results show that discrimination was readily achieved after the first washing step with no need to increase the stringency of the washes (Figure 3.7). Further washes with urea, showed to have a negligible effect on the non-complementary hybridisation. However, the complementary duplex strongly denatured with increasing stringency.

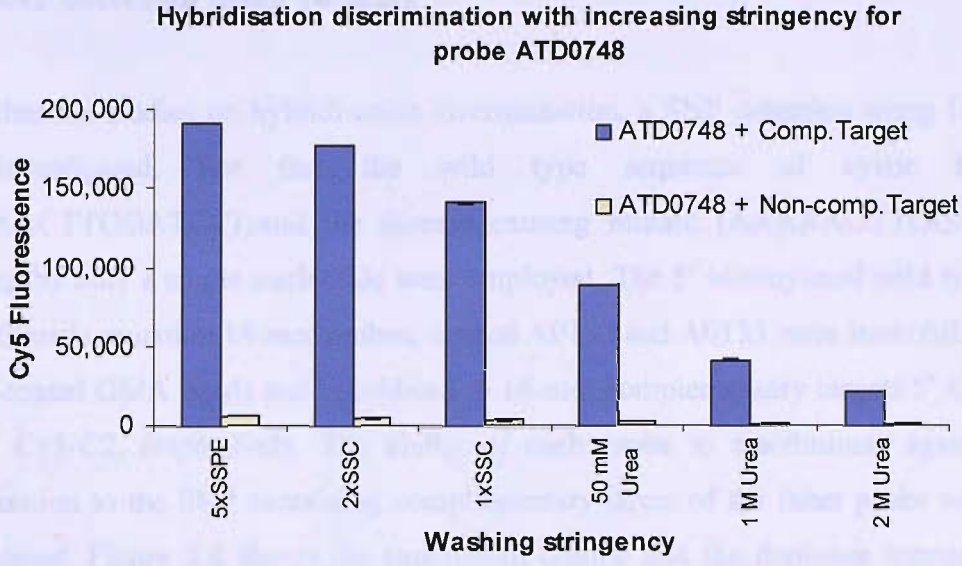


Figure 3.7 Discrimination between complementary and non-complementary hybridisation for probe ATD0748.

The discrimination in hybridisation exhibited by probe ATD0748 was expected due to the number of mismatched bases highly destabilising the non-complementary duplex. The reason why probe ATD0747 failed to exhibit a similar level of discrimination between complementary and non-complementary hybridisation was surprising and it was not understood. These findings were consistent across experiments and were observed at 37, 45 and 55 °C (data not shown).

3.3.3 SNP detection using 14-mers

To further the studies on hybridisation discrimination, a SNP detection using 14-mers was investigated. For that the wild type sequence of cystic fibrosis (AAAAACTTGGATCC) and the disease causing mutant (AAAAAGTTGGATCC) differing by only a single nucleotide were employed. The 5' biotinylated wild type and cystic fibrosis mutation 14-mer probes, termed A0132 and A0133 were immobilised on avidin coated GMA beads and hybridised to 16-mer complementary targets 5' Cy5-C1 and 5' Cy3-C2, respectively. The ability of each probe to discriminate against the hybridisation to the SNP containing complementary target of the other probe was also investigated. Figure 3.8 shows the experiment outline and the duplexes formed upon hybridisation.

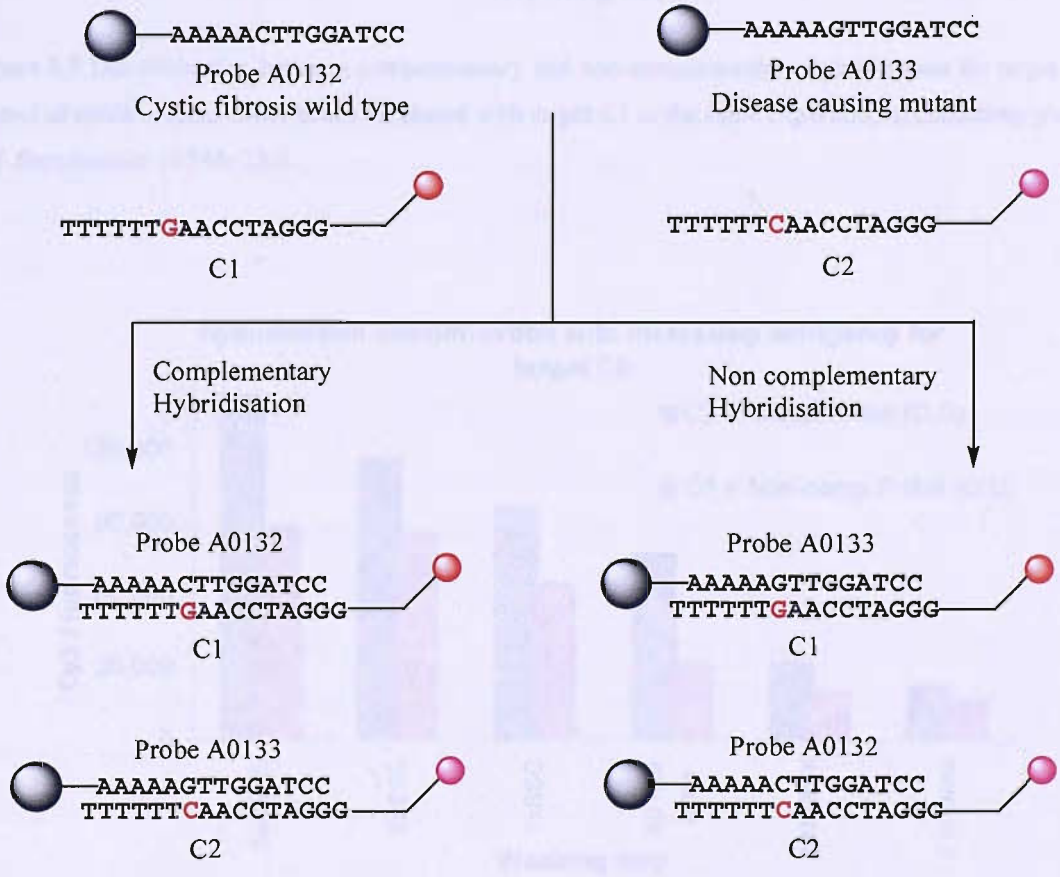


Figure 3.8 Outline of the duplexes formation for the cystic fibrosis wild type and disease causing mutant probes upon hybridisation. The 5' biotinylated probes were immobilised on avidin coated GMA beads.

All suspensions were incubated overnight at room temperature. The beads were then washed with buffers and urea solutions of increased stringency and analysed by FACS at each step (Figures 3.9 and 3.10).

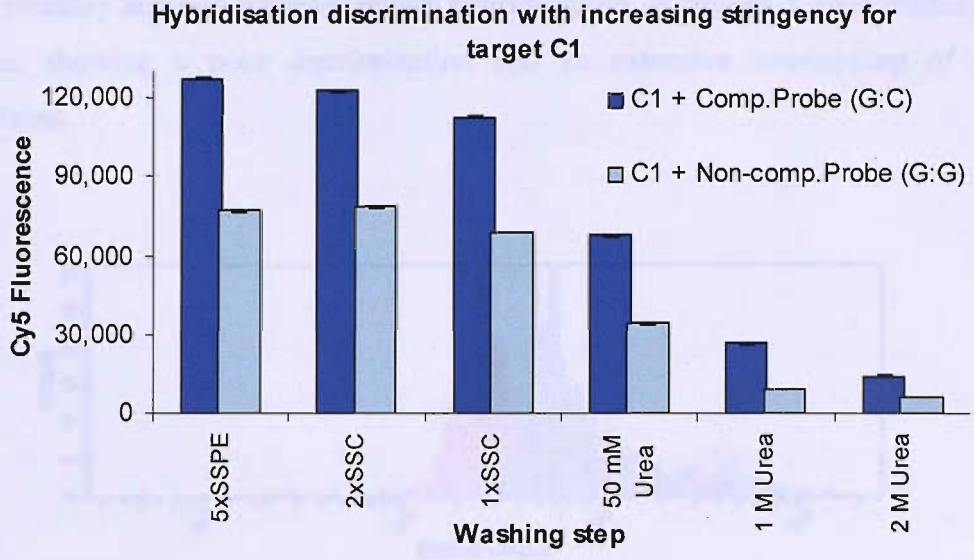


Figure 3.9 Discrimination between complementary and non-complementary hybridisation for target C1. A control of avidin coated GMA beads incubated with target C1 in the same experimental conditions yielded a Cy5 fluorescence of 746±23.6.

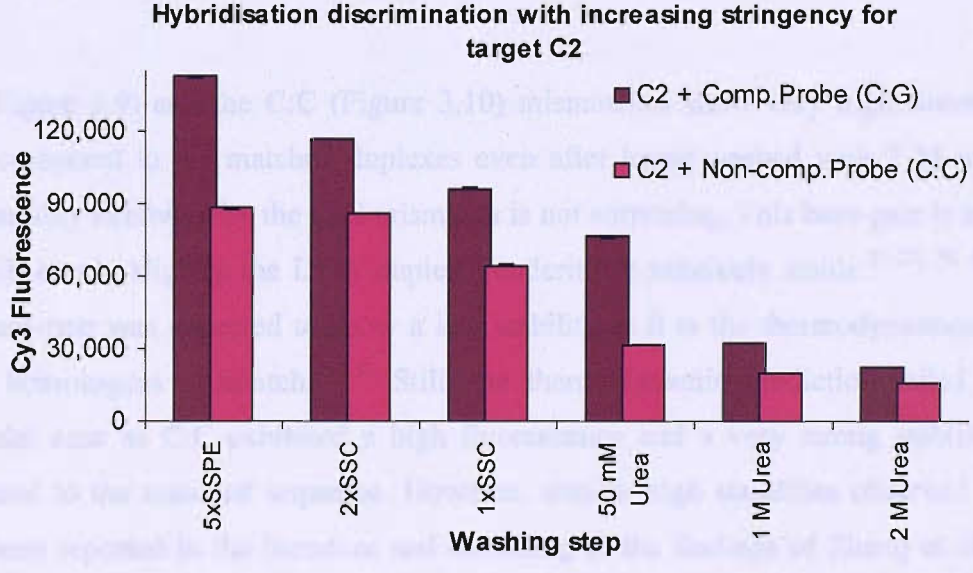


Figure 3.10 Discrimination between complementary and non-complementary hybridisation for target C2. A control of avidin coated GMA beads incubated with target C2 in the same experimental conditions yielded a Cy3 fluorescence of 107±0.6.

Figures 3.9 and 3.10 show that a difference between complementary and non-complementary hybridisation for both targets is readily observed from the outset. However, poor discrimination was observed throughout the washing steps. To illustrate, Figure 3.11 shows the Cy5 fluorescence profile for the mixture of the individual complementary and non-complementary hybridisations of target C1 after washing with 2 M urea, showing a poor discrimination and an extensive overlapping of the two populations.

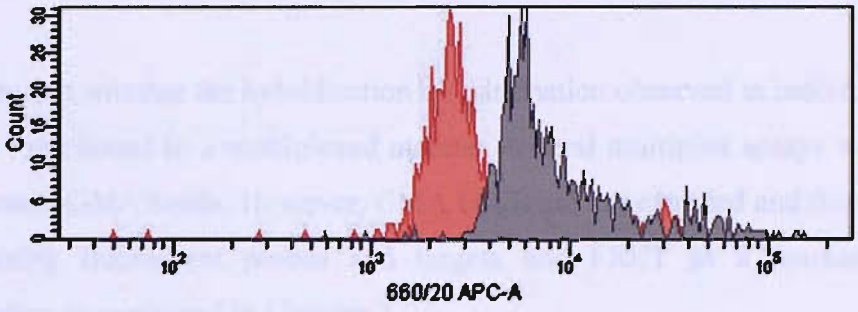


Figure 3.11 Cy5 fluorescence profile (detected by filter 660/20 APC-A) for the mixture of the individual complementary and non-complementary hybridisations of target C1 after washing with 2 M urea. Grey and red profiles represent complementary and mismatched populations respectively. Count refers to the number of recorded events (particles detected).

G:G (Figure 3.9) and the C:C (Figure 3.10) mismatches show very high fluorescences when compared to the matched duplexes even after being washed with 2 M urea. The high stability exhibited by the G:G mismatch is not surprising. This base-pair is known to destabilise only slightly the DNA duplex, rendering it relatively stable.[25-27, 30, 44-46] The C:C base-pair was expected to show a low stability as it is the thermodynamically least stable homologous mismatch.[31, 32] Still, the thermodynamic prediction failed for this particular case as C:C exhibited a high fluorescence and a very strong stability when compared to the matched sequence. However, similar high stabilities observed for C:C have been reported in the literature and according to the findings of Zhang et al. C:C is generally a stable mismatch in the context of microarrays.[19, 31, 44] This mismatched stability and poor single base discrimination was consistent throughout experiments and verified at 34, 40 and 45 °C (data not shown).

3.3.4 Multiplex assays on avidin coated GMA beads

Multiplexing bead-based suspension technologies, which allow for simultaneous detection of multiple target sequences in a single reaction, can seriously reduce the time, cost and labour associated with single reaction detection technologies.[47] Optimising hybridisation efficiency and discrimination is paramount for bead-based arrays, but also low levels of cross hybridisation are important for assessing and quantifying abundances of DNA targets in environmental samples containing many unknown sequences and in different concentrations.

In order to test whether the hybridisation discrimination observed in individual reactions could be reproduced in a multiplexed manner, several multiplex assays were set up on avidin coated GMA beads. However, GMA beads are not encoded and this problem was solved using fluorescent probes and targets and FRET as a marker for correct hybridisation as explained in Chapter 2.

3.3.4.1 Assessing the cross-reactivity of two immobilised probes and one target in solution

To assess the cross-reactivity of two completely unrelated immobilised probes and one target in solution, 5' biotinylated - 3' fluorescein 20-mer (ATD1200) and 5' biotinylated 26-mer (ATD1198) probes were immobilised on GMA beads and were incubated at room temperature for fifteen minutes with target, 5' Cy5 20-mer, ATD960 complementary to probe ATD1200 (Figure 3.12). The levels of complementary and non-complementary hybridisation were assessed by individual incubations of the immobilised probes with the target. The beads were washed with buffers of increased stringency and analysed by FACS.

Figure 3.12 outlines the multiplex experiment and the FRET generated by the complementary hybridisation of ATD1200 and ATD0960 sequences, due to the presence of the fluorescein and Cy5 fluorophores.

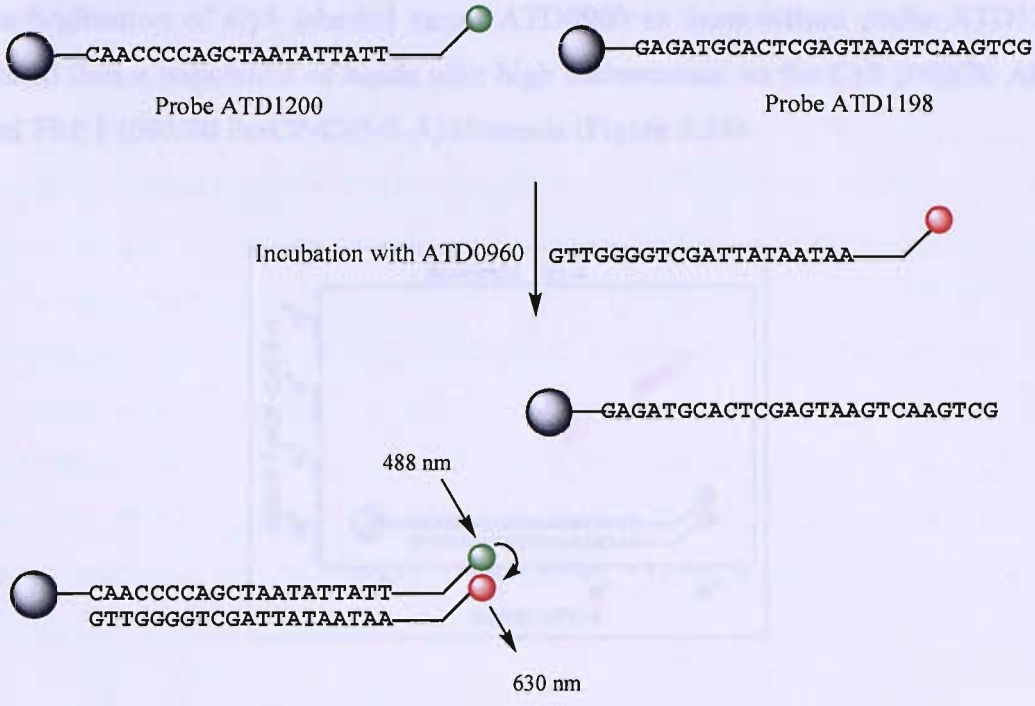


Figure 3.12 Cross-reactivity hybridisation experiment outline. The hybridisation of probe ATD1200 and complement ATD0960 generates FRET due to the close proximity of fluorescein and Cy5 fluorophores.

The immobilisation of probe ATD1200, resulted in a population of beads highly fluorescent on the fluorescein channel (530/30 FITC-A) (Figure 3.13A). Upon hybridisation to complementary target ATD0960, fluorescein fluorescence was drastically reduced (Figure 3.13B), suggesting a channelling of energy transfer between fluorophores via FRET.

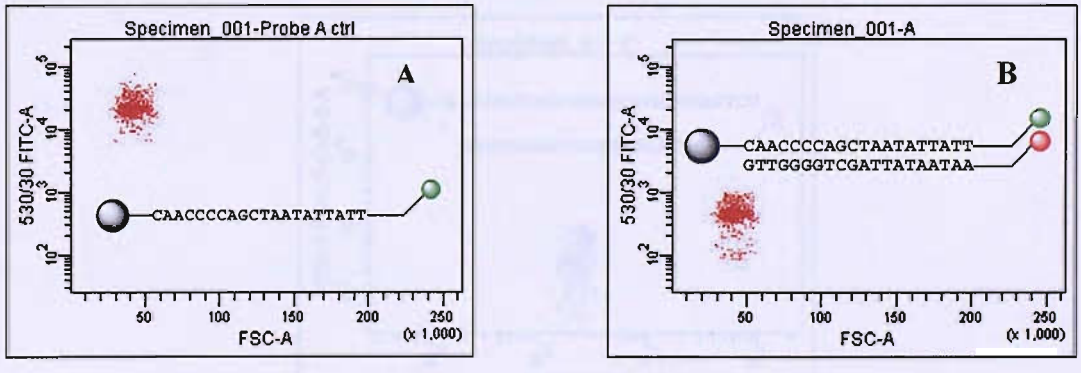


Figure 3.13 FACS scatter plot - Fluorescein fluorescence (detected by filter 530/30 FITC-A) of immobilised probe ATD1200, before (A) and after (B) hybridisation to Cy5 labelled target. FSC-A (Forward scatter) provides information about the size of the analysed microparticles.

The hybridisation of Cy5 labelled target ATD0960 to immobilised probe ATD1200 generated thus a population of beads with high fluorescence on the Cy5 (660/20 APC-A) and FRET (695/40 PerCP-Cy5-5-A) channels (Figure 3.14).

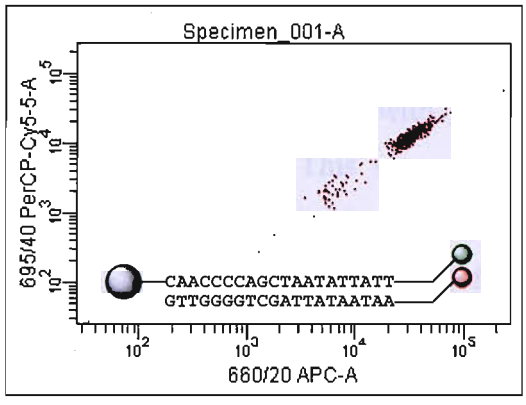


Figure 3.14 Cy5 and FRET fluorescence for the complementary hybridisation of probe ATD1200 to target A0960. 660/20 APC-A and 695/40 PerCP-Cy5-5-A are the filters used to detect Cy5 and FRET fluorescence, respectively.

Probe ATD1198 exhibited some non-specific binding to non-complementary target ATD0960 (Figure 3.15). No fluorescein or FRET fluorescence was observed, as expected, as the probe was non-labelled.

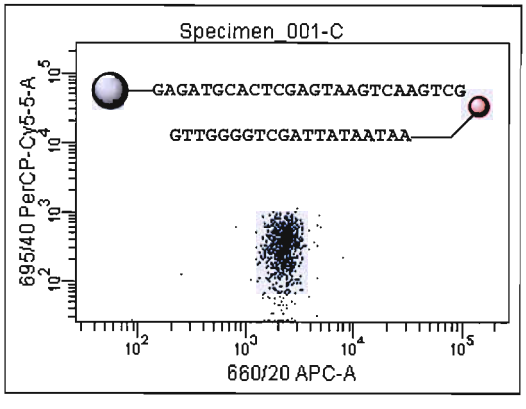


Figure 3.15 Cy5 and FRET fluorescence for the non-complementary hybridisation of probe ATD1198 to target A0960.

The multiplex assay, involving both immobilised probes and target generated two different fluorescent populations as expected (Figure 3.16). One population (red) exhibited high levels of Cy5 and FRET fluorescence indicating the complementary hybridisation between probe ATD1200 and target ATD0960. A second population (blue) exhibited low levels of Cy5 and FRET fluorescence indicating the non-complementary hybridisation between probe ATD1198 and ATD0960. The multiplex assay displayed in Figure 3.16 correlated well with the individual hybridisation reactions showed in Figures 3.14 and 3.15. This suggested very good hybridisation discrimination between the species involved in the multiplex assay. In fact, the blue population in the multiplex assay showed a decrease in Cy5 fluorescence. This indicated an improvement on the discrimination against the non-complementary hybridisation between ATD1198 and ATD0960.

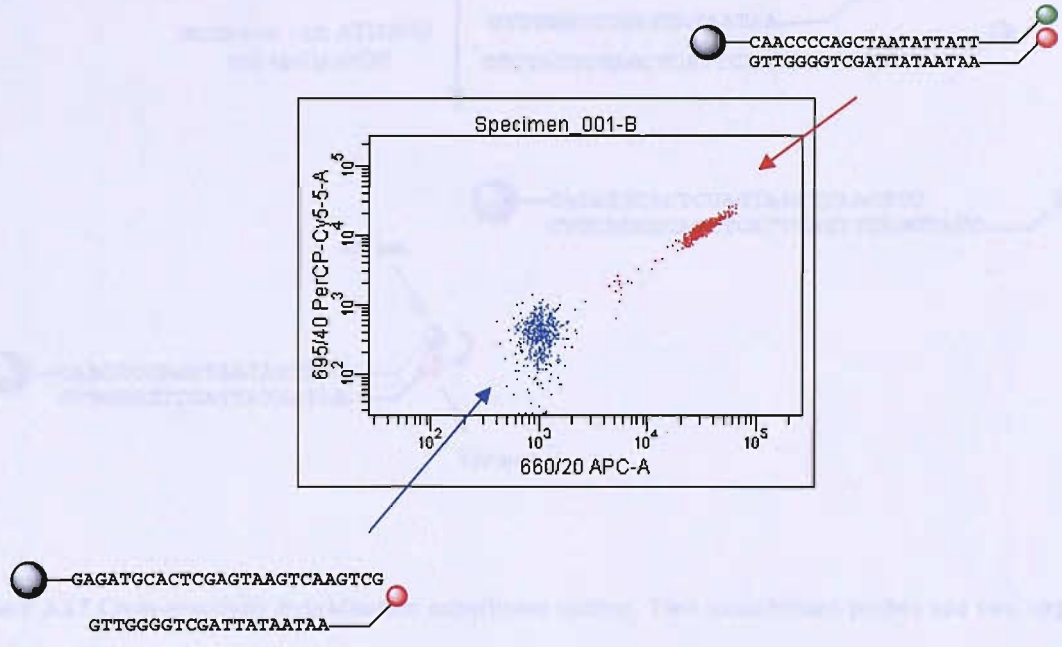


Figure 3.16 Multiplex assay showing two different fluorescent populations.

3.3.4.2 Assessing the cross-reactivity of two immobilised probes and two targets in solution

To assess the cross-reactivity of two completely unrelated immobilised probes and two targets, a 5' fluorescein 30-mer target (4goligo0020), complementary to probe ATD1198 was added to the system described in the previous section (Figure 3.17). The experiment was carried out as described before and the results are displayed in Figure 3.18.

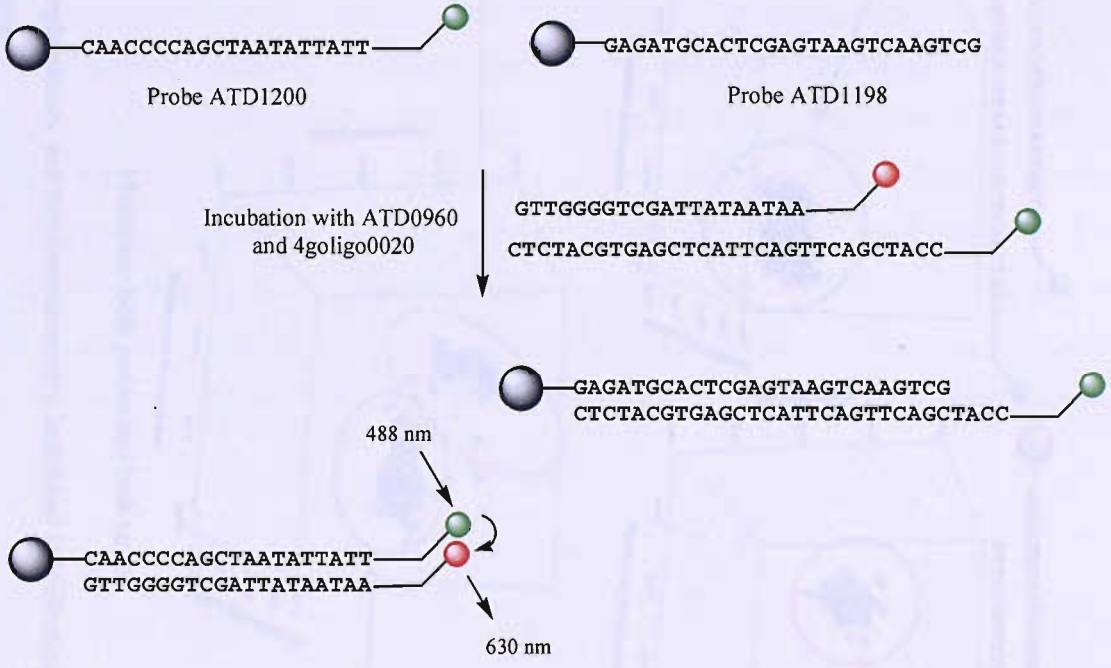


Figure 3.17 Cross-reactivity hybridisation experiment outline. Two immobilised probes and two targets in solution.

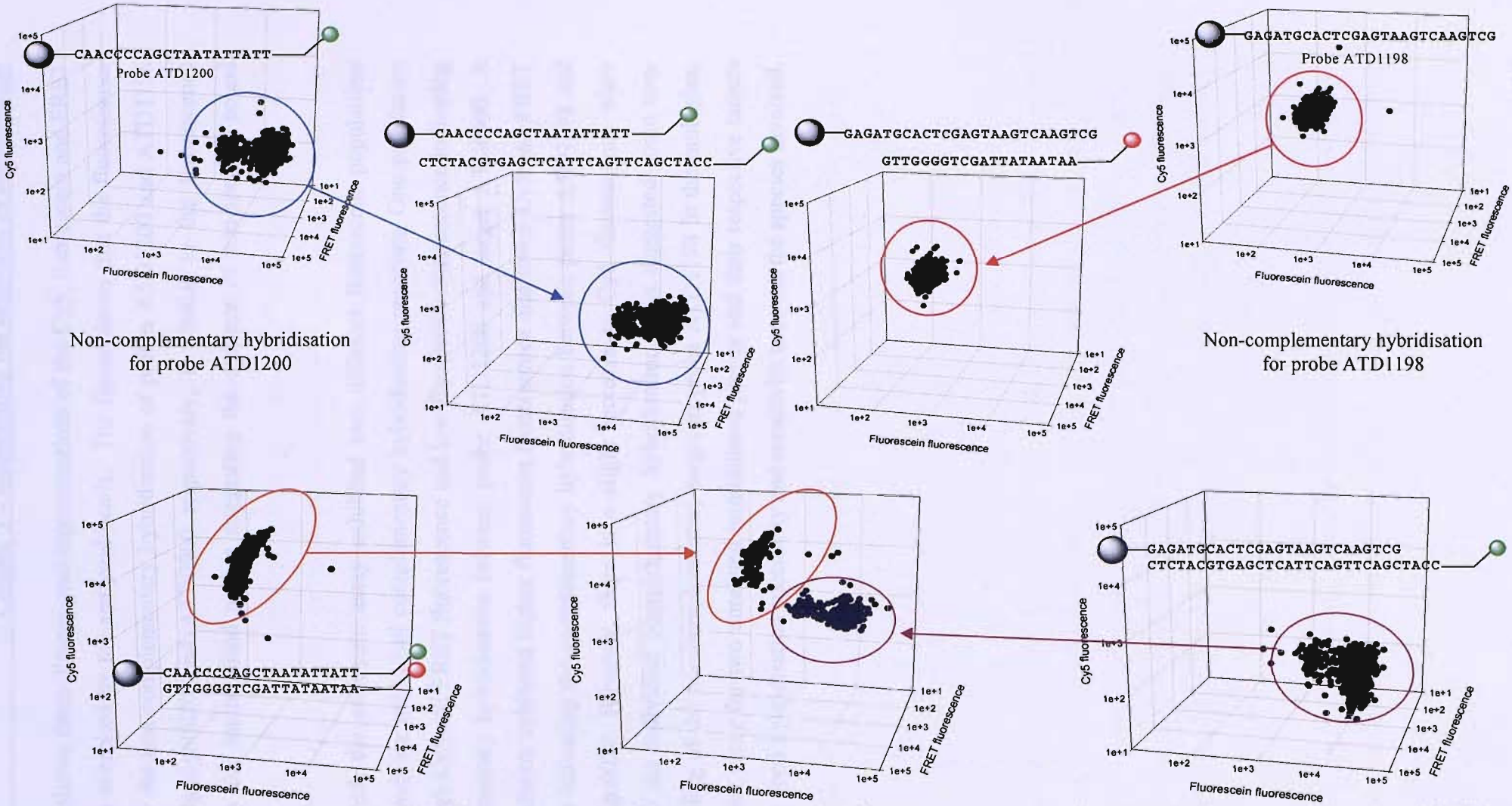


Figure 3.18 Fluorescence of the beads for the complementary and non-complementary individual hybridisation reactions and for the multiplex assay. Three-dimensional plots showing Cy5, fluorescein and FRET fluorescence.

Figure 3.18 shows three-dimensional representations of the Cy5, fluorescein and FRET fluorescence exhibited by the analysed beads. The figure shows that the fluorescence exhibited by the non-complementary hybridisation of probes ATD1200 and ATD1198 to targets 4goligo0020 and ATD0960 respectively, is similar to the fluorescence exhibited by the immobilised probes suggesting the absence of non-specific bound targets.

As anticipated, the multiplex assay exhibited two different fluorescent populations correlating well the individual complementary hybridisation reactions. One population exhibited high Cy5 and FRET fluorescence and low fluorescein fluorescence indicating the complementary hybridisation between probe ATD1200 and target ATD0960. A second population exhibited higher fluorescein fluorescence and lower Cy5 and FRET fluorescence indicating the complementary hybridisation between probe ATD1198 and target 4goligo0020. However, there is a slight increase in Cy5 fluorescence when compared to the individual complementary hybridisation. This suggested some non-specific binding of the non-complementary target to probe ATD1198 in the multiplex. This multiplex assay for two unrelated immobilised probes and their respective targets yielded very good hybridisation specificity and sensitivity between the species involved.

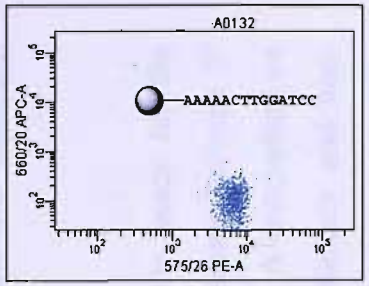
3.3.4.3 Assessing the cross-reactivity of two immobilised probes differing in one base and two targets in solution

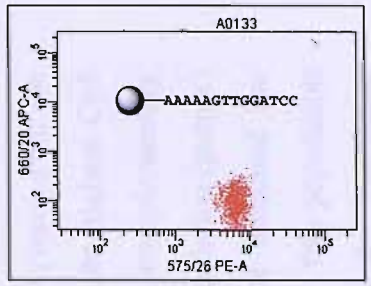
Previous sections dealt with multiplex assays involving two completely unrelated probes. This section investigates a multiplex assay involving two immobilised probes differing in one base. The system considered in section 3.3.3, for the SNP detection using the wild type sequence of cystic fibrosis and the disease causing mutant was further studied to investigate whether the single nucleotide discrimination would be the same in a multiplex assay. Figure 3.19 reviews the duplexes formed upon complementary hybridisation.

Figure 3.19 Wild type sequence of cystic fibrosis and the disease causing mutant and their respective complementary targets.

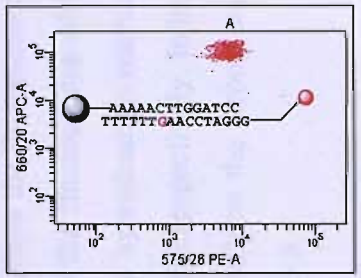
The reactions to assess the level of complementary and non-complementary hybridisation between the targets and the immobilised probes were carried out as described in section 3.3.3. Multiplexed assays consisted in the incubation of both probes with 5' Cy5-C1 target, both probes with 5' Cy3-C2 target and both probes with both targets. Incubation was carried out at room temperature for fifteen minutes. The beads were then washed with buffers and urea solutions of increased stringency and analysed by FACS.

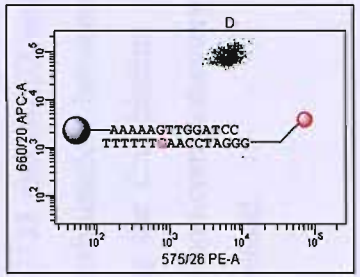
Figure 3.20 shows the fluorescence of the beads, on the Cy5 (660/20 APC-A) and Cy3 (575/26 PE-A) channels, for the complementary and non-complementary hybridisation reactions and the multiplex, after being washed with buffers.



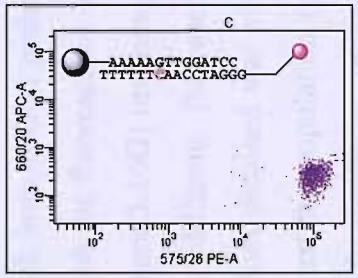


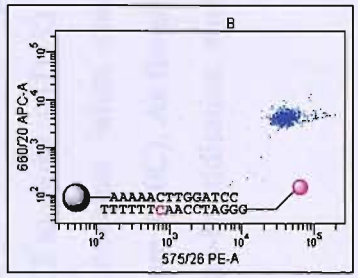
Complementary and non-complementary hybridisation for target C1



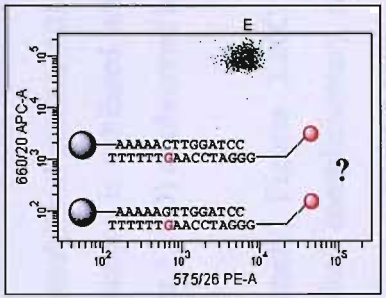


Complementary and non-complementary hybridisation for target C2

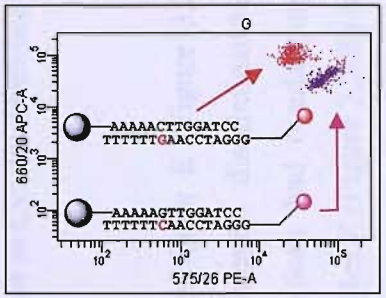




Multiplex: Both probes and target C1



Multiplex: Both probes and both targets



Multiplex: Both probes and target C2

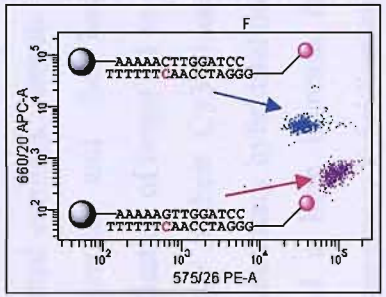


Figure 3.20 Fluorescence of the beads for the complementary and non-complementary individual hybridisation reactions between targets and immobilised probes and for the multiplex assays, after being washed with buffers of increased stringency. 575/26 PE-A and 660/20 APC-A are the filters used to detect Cy3 and Cy5 fluorescence, respectively.

Figures 3.20 (A0132) and 3.20 (A0133) show the fluorescence profile for the immobilised probes. Figures 3.20A and 3.20D show the fluorescence of the beads for the individual complementary and non-complementary hybridisation reactions between 5' Cy5-C1 and the immobilised probes, A0132 and A0133 respectively. In both cases a high level of Cy5 fluorescence is exhibited which again reflects the strong stability of the G:G mismatch (Figure 3.16D) as observed in section 3.3.3.

Conversely, Figures 3.20C and 3.20B show the fluorescence of the beads for the individual complementary and non-complementary hybridisation reactions between 5' Cy3-C2 and the immobilised probes, A0133 and A0132 respectively. A comparison on the Cy3 channel revealed a poor discrimination between complementary and non-complementary hybridisation, illustrating the C:C binding stability as observed before. The fact that C2 upon hybridisation to non-complementary probe A0132 exhibited an increase in Cy5 fluorescence is not understood.

For multiplex E (Figure 3.20E) involving both immobilised probes and target C1, hybridisation discrimination was not possible as both matched and mismatched populations had overlapping fluorescence and only one population was observed. Multiplex F (Figure 3.20F) involving both immobilised probes and target C2 generated two fluorescent populations consistent with the fluorescence intensities of the individual hybridisation controls. Multiplex G (Figure 3.16G) involving both immobilised probes and both targets also generated two different fluorescent populations, but the interpretation of the fluorescence intensities obtained was not straightforward from the individual controls. Results suggested that the population (red) with the highest Cy5 fluorescence and the lowest Cy3 fluorescence indicated the complementary hybridisation of target C1 to probe ATD132. On the other hand, the population (purple) with the highest Cy3 fluorescence and the lowest Cy5 fluorescence indicated the complementary hybridisation of target C2 to probe ATD133. However, this population showed a high increase in Cy5 fluorescence when compared to the individual complementary hybridisation of C2 (Figure 3.20C). As the target is Cy3 labelled, this is not understood. The high levels of cross-hybridisation must not be ignored and are certainly influencing the results.

Washing the beads further with urea did not make an improvement on the hybridisation discrimination which is consistent with that observed in section 3.3.3. Both for targets C1 and C2 the level of complementary and non-complementary overlapping observed made population discrimination in the multiplexes impossible (Figure 3.21).

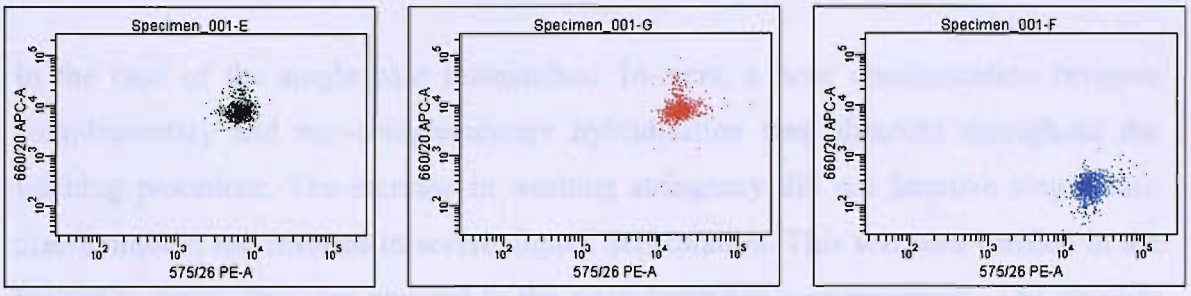


Figure 3.21 Fluorescence of the multiplex assays after washing with 2 M urea. E – Multiplex: Both probes and target C1. G – Multiplex: Both probes and both targets. F – Multiplex: Both probes and target C2.

The low hybridisation specificity and low sensitivity obtained for the multiplex of the single mutation 14-mers is consistent with that shown in section 3.3.3. The level of cross-hybridisation observed was the same.

This system demonstrates the difficulty in hybridisation discrimination between complementary and non-complementary duplexes for closely related sequences in a multiplex assay using short probes. This emphasises the need to carefully study and construct probes in order to maximise complementary over non-complementary hybridisation, avoiding situations that can compromise an array platform.

3.4 Conclusions

Experimental results showed that probes can each discriminate against the hybridisation to completely unrelated targets. This was true for the individual hybridisation reactions and for the multiplex assays. Washing the beads with buffers of increased stringency was sufficient to attain hybridisation discrimination.

In the case of the single base mismatched 14-mers, a poor discrimination between complementary and non-complementary hybridisation was observed throughout the washing procedure. The increase in washing stringency did not improve single base discrimination and resulted in severe duplex denaturation. This was also verified in the multiplex assay. This was credited to the relatedness between sequences. The stability observed for the G:G mismatch was not surprising as this mismatched base-pair can be very stable. The high stability of the C:C mismatch was unexpected as this base-pair is the thermodynamically least stable homologous mismatch.[31, 32] However, according to the literature similar high stabilities have been reported and according to the findings of Zhang and co-workers C:C is generally a stable mismatch in the context of microarrays.[19, 31, 44]

The experimental results for the unrelated 16-mers were surprising. It was not understood why one probe discriminated strongly against non-complementary hybridisation and the other probe did not, when both probes, upon non-complementary hybridisation formed the same number of matched and mismatched base-pairs in the same position. The order of the mismatched base pairs in both duplexes was different, and that might affect the stacking of the base pairs in their vicinity leading to completely different stabilities and behaviour, explaining the results obtained. Also, a similar situation to that of the 14-mers was observed; the increase in washing stringency increased the poor discrimination.

The DNA/PNA 16-mers time course hybridisation experimental results suggested that fifteen minutes of incubation as opposed to an overnight incubation, resulted in lower non-specific binding of non-complementary targets. However, the experiments carried

out with the single base mismatched 14-mers (sections 3.3.3 and 3.3.4.3), overnight and for fifteen minutes generated similar results.

Throughout the course of optimising the experimental conditions for better hybridisation selectivity, the experiments were performed at different temperatures. Results showed that room temperature incubations resulted in higher levels of fluorescence than incubations at higher temperatures. This is not surprising as the combination of a high salt concentration (5xSSPE, 0.1 % SDS) and low temperature incubation reduces the hybridisation stringency and enhances the stability of the duplexes.[20, 24]

The multiplex assays involving two completely unrelated probes and one or two targets were successfully carried out. These multiplexes exhibited different populations consistent with the individual hybridisation reactions. On the other hand, the multiplexes involving the 14-mers exhibited a poor hybridisation discrimination and consequently a poor population discrimination due to high stability of the single mismatches. The multiplex systems studied are a promising tool in that when applied to encoded solid supports, multiplexed assays will have unequivocal information relating the code and therefore the probe sequence to the fluorescence of the particle, without the need for so many individual fluorescence controls. Other work in our group, using the same 14-mers and encoded SU8 microparticles, provided clear results in that it was possible to verify and quantify the complementary and non-complementary hybridisation by analysing the fluorescence of the particles and their codes.

The combination of the experimental results obtained for the set of short probes studied is confusing. In general, closely related sequences resulted in low hybridisation specificity and low sensitivity. One immediate conclusion is that a good probe design strategy is essential for accurate expression measurement.[17, 20] So far, choosing optimal probes has been difficult because of the present poor understanding of how sensitivity and specificity of a probe depend upon its length and sequence. [9, 16] Although duplex formation in solution has been extensively studied using the nearest-neighbour model,[25-29, 31, 32, 48] hybridisation still involves a multitude of variables not yet predictable by computational methods.[17] Consequently, it has been difficult to transfer these findings and predictions to arrays. Binding interactions on arrays seem to be complicated by a

series of factors such as steric hindrance on the surface of the solid support and probe-probe interaction.[18]

Some of the results obtained in this chapter were surprising and were not understood. This prompted the study carried out in chapter 4. A comprehensive set of single base mismatches was investigated in order to provide some answers to the questions raised in this chapter.

3.5 Experimental Methods

3.5.1 Reagents

Amino-modified GMA beads (5 - 6 μm - 10% (W/V)) were obtained from Bangs Labs. Oligonucleotide sequences were synthesised by ATDbio and Sigma-Aldrich (Tables 3.2 and 3.3). Avidin DN (1.0 mg/ml) was from Vector Laboratories. EDC (1-ethyl-3-(3-dimethylaminopropyl) carbodiimide), sodium chloride, sodium citrate, sodium dihydrogen phosphate, EDTA, SDS (sodium dodecyl sulphate) and urea were obtained from Acros and imidazole came from Aldrich.

DNA Probes	Type of modification	Sequence 5' – 3'	Length (nt)
ATD0747	5' Biotin	TTGTTATAGTTCTCTC	16
ATD0748	5' Biotin	CTAGTTACTCTTGTTC	16
A0132	5' Biotin	AAAAACTTGGATCC	14
A0133	5' Biotin	AAAAAGTTGGATCC	14
PNA161	5' Biotin	CTAGTTACTCTTGTTC	16
ATD1198	5' Biotin	GAGATGCACTCGAGTAAGTCAAGTCG	26
ATD1200	5' Biotin 3' Fluorescein	CAACCCCAGCTAATATTATT	20

Table 3.2 DNA and PNA probe sequences employed.

DNA Targets	Type of modification	Sequence 5' – 3'	Length (nt)
ATD0960	5' Cy5	AATAATATTAGCTGGGGTTG	20
ATD0749	5' Cy5	GAGAGAACTATAACAA	16
ATD0750	5' Cy5	GAACAAGAGTAACTAG	16
C1	5' Cy5	GGGATCCAAGTTTTTT	16
C2	5' Cy3	GGGATCCAACTTTTTT	16
4goligo0020	5' Fluorescein	CCATCGACTTGACTTACTCGAGTGCATCTC	30
4goligo0016	Cy5	CCATCGACTTGACTTACTCGAGTGCATCTC	30

Table 3.3 DNA target sequences employed.

3.5.2 FACS analysis

The experimental methodology for the FACS analysis of the samples of beads was performed as described in section 2.5.2.

3.5.3 Preparation of avidin functionalised solid supports

The experimental methodology for the avidin functionalisation of GMA beads is described in section 2.5.14.

3.5.4 Covalent attachment of biotinylated DNA and PNA to avidin coated GMA beads

The experimental methodology for the covalent attachment of biotinylated DNA and PNA probes to avidin coated GMA beads was as described in section 2.5.15.

3.5.5 DNA hybridisation on avidin coated GMA beads and post-hybridisation stringency

The experimental methodology for DNA/PNA hybridisation follows the same procedure described in section 2.5.16. Hybridisation incubations were carried out at constant temperature with constant agitation either overnight or for fifteen minutes. The

DNA/PNA time course experiment was preformed in the same way and aliquots were taken at the designated times for monitoring the reaction progress. After hybridisation the beads were washed with buffers (5xSSPE, 0.1% SDS, 2xSSC, 0.1% SDS, 1xSSC, 0.1% SDS) and/or urea solutions (50 mM, 1 M, 2 M) of increased stringency and analysed by FACS.

3.5.6 Multiplexed assays

DNA probes were immobilised and hybridised to targets as described. Multiplex assays were performed as described for individual assays and contained the same volume and concentration of GMA immobilised probes. Multiplex hybridisation to more than one target in solution contained a tenfold excess of each target. Hybridisation conditions involved the mixture of the required participants in 200 μL of hybridisation buffer and reactions were agitated for fifteen minutes at room temperature, after which the beads were washed with buffers and/or urea solutions of increased stringency.

3.6 References

1. Dai, H.; Meyer, M.; Stepaniants, S.; Ziman, M.; Stoughton, R., Use of hybridization kinetics for differentiating specific from non-specific binding to oligonucleotide microarrays. *Nucleic Acids Res.* **2002**, 30, e86.
2. Lemeshko, S.V.; Powdrill, T.; Belosludtsev, Y.Y.; Hogan, M., Oligonucleotides form a duplex with non-helical properties on a positively charged surface. *Nucleic Acids Res.* **2001**, 29, 3051–8.
3. Penchovsky, R.; Birch-Hirschfeld, E.; McCaskill, J.S., End-specific covalent photo-dependent immobilisation of synthetic DNA to paramagnetic beads. *Nucleic Acids Res.* **2000**, 28, E98.
4. Bunemann, H.; Westhoff, P.; Herrmann, R.G., Immobilization of denatured DNA to macroporous supports: I. Efficiency of different coupling procedures. *Nucleic Acids Res.* **1982**, 10, 7163-80.
5. Bunemann, H., Immobilization of denatured DNA to macroporous supports: II. Steric and kinetic parameters of heterogeneous hybridization reactions *Nucleic Acids Res.* **1982**, 10, 7181-96.
6. Dorris, D.R.; Nguyen, A.; Gieser, L.; Lockner, R.; Lublinsky, A.; Patterson, M.; Touma, E.; Sendera, T.J.; Elghanian, R.; Mazumder, A., Oligodeoxyribonucleotide probe accessibility on a three-dimensional DNA microarray surface and the effect of hybridization time on the accuracy of expression ratios. *BMC Biotechnol.* **2003**, 3, 6-17.

7. Chan, V.; Graves, D.J.; McKenzie, S.E., The biophysics of DNA hybridization with immobilized oligonucleotide probes. *Biophys. J.* **1995**, 69, 2243-55.
8. Maskos, U.; Southern, E.M., Oligonucleotide hybridizations on glass supports: a novel linker for oligonucleotide synthesis and hybridization properties of oligonucleotides synthesised in situ. *Nucleic Acids Res.* **1992**, 7, 1679-84.
9. Southern, E.M.; Mir, K.; Shchepinov, M., Molecular interactions on microarrays. *Nat. Genet.* **1999**, 21, 5-9.
10. Bhanot, G.; Louzoun, Y.; Zhu, J.; DeLisi, C., The importance of thermodynamic equilibrium for high throughput gene expression arrays. *Biophys. J.* **2003**, 1, 124-35.
11. Melchior, W.B.; Von-Hippel, P.H., Alteration of the relative stability of dA-dT and dG-dC base pairs in DNA. *Proc. Natl. Acad. Sci. USA* **1973**, 2, 298-302.
12. Zheng, D.; Alm, E.W.; Stahl, D.A.; Raskin, L., Characterization of universal small-subunit rRNA hybridization probes for quantitative molecular microbial ecology studies. *Appl. Environ. Microbiol.* **1996**, 12, 4504-13.
13. Casey, J.; Davidson, N., Rates of formation and thermal stabilities of RNA:DNA and DNA:DNA duplexes at high concentrations of formamide. *Nucleic Acids Res.* **1977**, 5, 1539-52.
14. Sambrook, J.; Fritsch, E.F.; Maniatis, T., Molecular Cloning - A Laboratory Manual. **1989**, Cold Spring Harbor Press, Cold Spring Harbor, New York (2.60-2.80), (2nd edn.).
15. Bates, S.R.; Baldwin, D.A.; Channing, A.; Gifford, L.K.; Hsu, A.; Lu, P., Cooperativity of paired oligonucleotide probes for microarray hybridization assays. *Anal. Biochem.* **2005**, 342, 59-68.
16. Doktycz, M.J.; Morris, M.; Dormady, S.J.; Beattie, K.L.; Jacobson, K.B., Optical melting of 128 octamer DNA duplexes. Effects of base pair location and nearest neighbors on thermal stability. *J. Biol. Chem.* **1995**, 15, 8439-45.
17. Chou, C.C.; Chen, C.H.; Lee, T.T.; Peck, K., Optimization of probe length and the number of probes per gene for optimal microarray analysis of gene expression. *Nucleic Acids Res.* **2004**, 32, e99.
18. Zhang, L.; Miles, M.; Aldape, K.D., A model of molecular interactions on short oligonucleotide microarrays. *Nat. Biotechnol.* **2003**, 21, 818-21.
19. Wu, C.; Carta, R.; Zhang, L., Sequence dependence of cross-hybridization on short oligo microarrays. *Nucleic Acids Res.* **2005**, 33, e84.
20. Strachan, T.; Read A.P., Human Molecular Genetics 2. *Bios* **1999**, John Wiley and Sons Inc., New York, (2ed).
21. Jayaraman, A.; Hall, C.K.; Genzer, J., Computer simulation study of molecular recognition in model DNA microarrays. *Biophys. J.* **2006**, 91, 2227-36.
22. Haslam, N., Optimising probe library design for DNA resequencing. *Ph.D. - University of Southampton* **2007**.
23. Ali, M.F.; Kirby, R.; Goodey, A.P.; Rodriguez, M.D.; Ellington, A.D.; Neikirk, D.P.; McDevitt, J.T., DNA hybridization and discrimination of single-nucleotide mismatches using chip-based microbead arrays. *Anal. Chem.* **2003**, 75, 4732-39.
24. Meinkoth, J.; Wahl, G., Hybridization of nucleic acids immobilized on solid supports. *Anal. Biochem.* **1984**, 138, 267-84.
25. Allawi, H.T.; SantaLucia, J.Jr., Thermodynamics and NMR of internal G.T mismatches in DNA. *Biochemistry* **1997**, 36, 10581-94.
26. Allawi, H.T.; SantaLucia, J.Jr., NMR solution structure of a DNA dodecamer containing single G.T mismatches. *Nucleic Acids Res.* **1998**, 21, 4925-34.
27. Allawi, H.T.; SantaLucia, J.Jr., Nearest neighbor thermodynamic parameters for internal G.A mismatches in DNA. *Biochemistry* **1998**, 37, (8), 2170-9.

28. Allawi, H.T.; SantaLucia, J.Jr., Thermodynamics of internal C.T mismatches in DNA. *Nucleic Acids Res.* **1998**, 26, 2694-701.
29. Allawi, H.T.; SantaLucia, J.Jr., Nearest-neighbor thermodynamics of internal A.C mismatches in DNA: sequence dependence and pH effects. *Biochemistry* **1998**, 37, 9435-44.
30. Ikuta, S.; Takagi, K.; Wallace, R.B.; Itakura, K., Dissociation kinetics of 19 base paired oligo-DNA duplexes containing different single mismatched base pairs. *Nucleic Acids Res.* **1987**, 15, 797-811.
31. Peyret, N.; Seneviratne, P.; Allawi, H.T.; SantaLucia, J.Jr., Nearest-neighbor thermodynamics and NMR of DNA sequences with internal A.A, C.C, G.G, and T.T mismatches. *Biochemistry* **1999**, 12, 3468-77.
32. Tikhomirova, A.; Beletskaya, I.; Chalikian, T.V., Stability of DNA duplexes containing GG, CC, AA, and TT mismatches. *Biochemistry* **2006**, 45, 10563-71.
33. Gingeras, T.R.; Kwoh, D.Y.; Davis, G.R., Hybridisation properties of immobilized nucleic acids. *Nucleic Acids Res.* **1987**, 15, 5373-90.
34. McQuain, M.K.; Seale, K.; Peek, J.; Fisher, T.S.; Levy, S.; Stremler, M.A.; Haselton, F.R., Chaotic mixer improves microarray hybridization. *Anal. Biochem.* **2004**, 15, 215-26.
35. Schaupp, C.J.; Jiang, G.; Myers, T.G.; Wilson, M.A., Active mixing during hybridization improves the accuracy and reproducibility of microarray results. *Biotechniques* **2005**, 38, 117-9.
36. Wilson, R.; Cossins, A.R.; Spiller, D.G., Encoded microcarriers for high-throughput multiplexed detection. *Angew. Chem. Int. Ed. Engl.* **2006**, 45, 6104-17.
37. Ratilainen, T.; Holmen, A.; Tuite, E.; Nielsen, P.E.; Norden, B., Thermodynamics of sequence-specific binding of PNA to DNA. *Biochemistry* **2000**, 26, 7781-91.
38. Ratilainen, T.; Lincoln, P.; Norden, B., A simple model for gene targeting. *Biophys. J.* **2001**, 5, 2876-85.
39. http://www.chemie.hu-berlin.de/seitz/research.htm.
40. Ratilainen, T.; Holmen, A.; Tuite, E.; Haaima, G.; Christensen, L.; Nielsen, P.E.; Norden, B., Hybridization of peptide nucleic acid. *Biochemistry* **1998**, 35, 12331-42.
41. Kambhampati, D.; Nielsen P.E.; Knoll, W., Investigating the kinetics of DNA-DNA and PNA-DNA interactions using surface plasmon resonance-enhanced fluorescence spectroscopy. *Biosens. Bioelectron.* **2001**, 9-12, 1109-18.
42. Jensen, K.K.; Orum, H.; Nielsen, P.E.; Norden, B., Kinetics for hybridization of peptide nucleic acids (PNA) with DNA and RNA studied with the BIAcore technique. *Biochemistry* **1997**, 16, 5072-7.
43. Mokhir, A.A.; Kraemer, R., Conjugates of PNA with naphthalene diimide derivatives having a broad range of DNA affinities. *Bioconjug. Chem.* **2003**, 14, 877-83.
44. Bhattacharya, P.K.; Cha, J.; Barton, J,K., 1H NMR determination of base-pair lifetimes in oligonucleotides containing single base mismatches. *Nucleic Acids Res.* **2002**, 30, 4740-50.
45. Hutton, JR., Renaturation kinetics and thermal stability of DNA in aqueous solutions of formamide and urea. *Nucleic Acids Res.* **1977**, 10, 3537-55.
46. Ke, S.H.; Wartell, R.M., Influence of nearest neighbor sequence on the stability of base pair mismatches in long DNA; determination by temperature-gradient gel electrophoresis. *Nucleic Acids Res.* **1993**, 22, 5137-43.
47. Dunbar, SA., Applications of Luminex xMAP technology for rapid, high-throughput multiplexed nucleic acid detection. *Clin. Chim. Acta* **2006**, 363, 71-82.

48. SantaLucia, J.Jr.; Allawi, H.T.; Seneviratne, P.A., Improved nearest-neighbor parameters for predicting DNA duplex stability. *Biochemistry* **1996**, 35, 3555-62.

4. Hybridisation selectivity for a comprehensive set of mismatches

4.1 Introduction

In addition to the Watson-Crick base pairs (G:C and T:A) there are eight possible single mismatches which occur in DNA with varying frequencies and stabilities. They are A:A, A:C, G:G, G:A, G:T, C:C, C:T, and T:T.[1] These mismatches have the potential to compromise the genetic stability of a cell. They occur as a result of base misincorporation during DNA replication, recombination, and from mutagenic chemicals and ionising radiation.[2, 3]

The methodologies for DNA analysis rely on the self-recognition properties of DNA and ability to discriminate against non-complementary hybridisation. The knowledge of thermodynamics of all possible base pairs, including single base mismatches, is an important step towards the development and design of better oligonucleotide probes and optimisation of experimental conditions.[2, 4] Thus, studies on structure and stability of mismatches are indispensable to minimise the chance of incorrect base pairing that can lead to the detection of wrong sequences.[1, 5]

The stability of mismatched duplexes depends on the number, identity and location of mismatches.[6] Mismatches located terminally are usually less destabilised to the duplex than those located internally.[1, 7] Also, the orientation adopted by the mismatched base-pair(s) influences the stability of the DNA duplex.[7]

4.1.1 Biophysics and thermodynamics of single base mismatches

4.1.1.1 Homologous mismatches

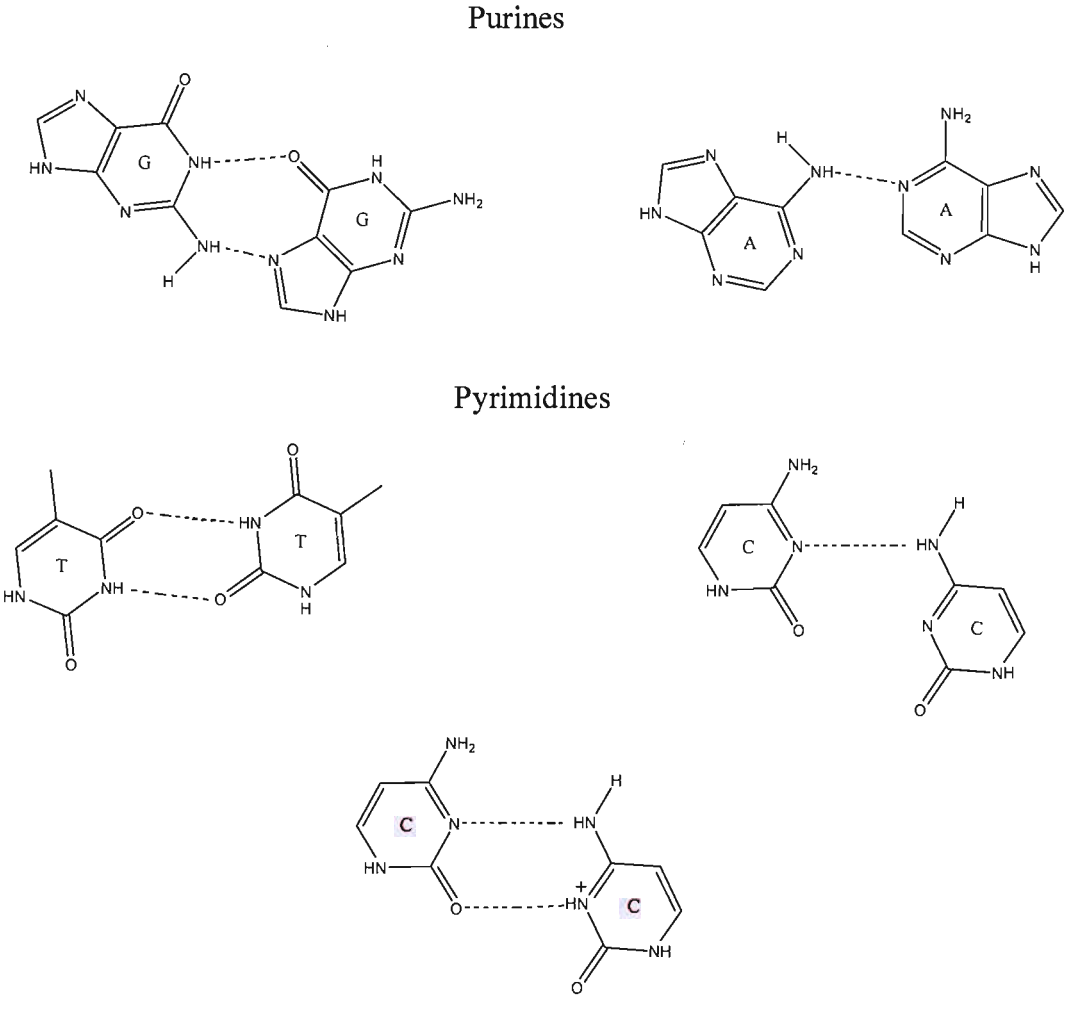


Figure 4.1 Structure of homologous mismatches; G:G, A:A, T:T, C:C and protonated C:C. Adapted from ref. 5.

Figure 4.1 shows the hydrogen bonds established by the homologous base-pair mismatches. The general stability trend for these mismatches published in the literature[4, 5, 8] is; G:G > T:T ≈ A:A > C:C. The relative thermodynamic stability for these base pairs can be rationalised in terms of the structure adopted by considering stacking energies and hydrogen bonding. Generally, G:G mismatches exhibit the highest thermodynamic stability. They have a high stacking potential due to their purine rings and are stabilised by two hydrogen bonds.[5, 8] Ranking next are T:T and A:A which

exhibit similar stability. T:T mismatches have a low stacking potential due to their pyrimidine rings and are stabilised by two hydrogen bonds in most contexts.[5] A:A, on the other hand have a higher stacking potential (Purine rings), but generally form one hydrogen bond. C:C are expected to be the most unstable mismatches of the homologous series as they have a low stacking potential and are stabilised by a single hydrogen bond.[5, 8] However, as seen in chapter 3, the C:C base pair was found to be a stable mismatch in the context of microarrays, which is in accordance with the findings of Zhang *et al.*[9] This suggests that care must be taken when extrapolating thermodynamic predictions evaluated from solution studies to the solid phase.

The stability trend has also been reported to depend on the surrounding sequence.[5, 8, 10] This context dependence has been seen to influence the stability of the G:G mismatch, for example. Depending on the surrounding base pairs, the G:G mismatch can exchange conformations and be stabilised by only one hydrogen bond.[11] The relative stabilities of A:A, T:T and C:C have also shown a context dependence, but no clear trend was observed.[5] Also, the thermodynamic impact of G:G, A:A and T:T is confined within the central triplet, whereas for C:C, the impact propagates into the adjacent helix domains and may involve up to nine base pairs.[8]

The pH can also influence the stability of mismatches. Peyret *et al.*[5] have reported an increase in stability for the C:C mismatch when lowering the pH from 7 to 4.9. This was probably due to the protonation of the cytosine rendering the base pair stabilised by two hydrogen bonds (Figure 4.1). Peyret also observed that the lowering in pH lowered the stability of A:A and T:T base pairs and did not affect greatly the G:G mismatch.[5]

The salt concentration can also affect the stability of the mismatches. Tikhomirova *et al.*[8] have reported an increase in T_m, and hence in stability of single base mismatches with increasing ionic strength of the solution. The T_m for the base pairs G:G, A:A and C:C changed linearly with increasing $[Na^+]$, while T:T exhibited a non-linear salt dependence.

4.1.1.2 Non-homologous mismatches

A:C mismatches

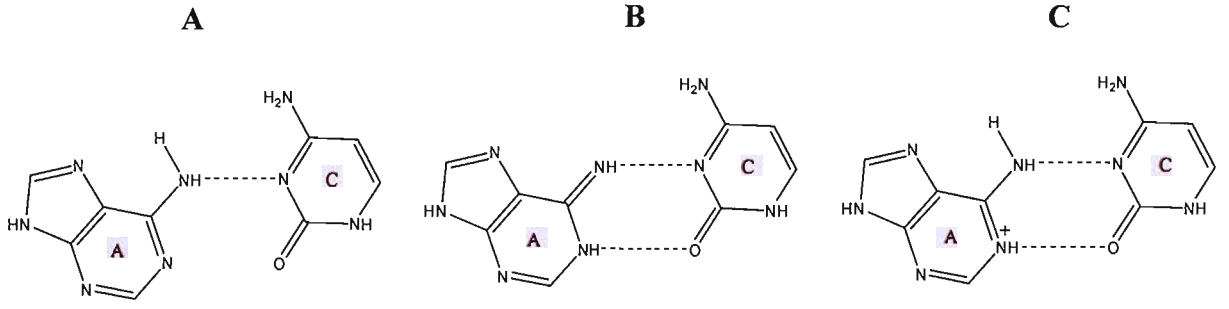


Figure 4.2 Structure of A:C mismatches at neutral pH (A and B) and at acidic pH (C). Adapted from ref. 2.

A:C mismatches have been shown to have a pH dependent thermodynamics and structure (Figure 4.2). At neutral pH, the A:C mismatch can exist in two configurations. It can be either stabilised by a single hydrogen bond as shown in Figure 4.2A or it can exist in a low population in the rare imino form, which stabilises the base-pair with two hydrogen bonds (Figure 4.2B). At low pH, the adenine can be protonated and the mismatched pair is stabilised by two hydrogen bonds (Figure 4.2C). The thermodynamics of the A:C mismatch and its contribution to the stability of the duplex is highly dependent on the sequence context.[2]

G:T mismatches

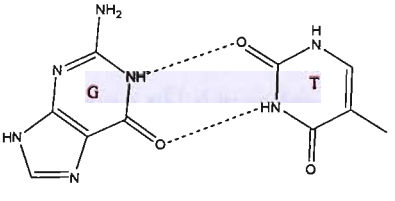


Figure 4.3 Structure of G:T mismatch. Adapted from ref. 6.

The single base mismatches containing guanine are usually the most thermodynamically stables.[10, 12] According to Allawi *et al.*[13] terminal G:T mismatches always make favourable contributions to the stability of the duplex, whereas internal G:T mismatches can make favourable or unfavourable contributions depending on the context sequence.

G:A mismatches

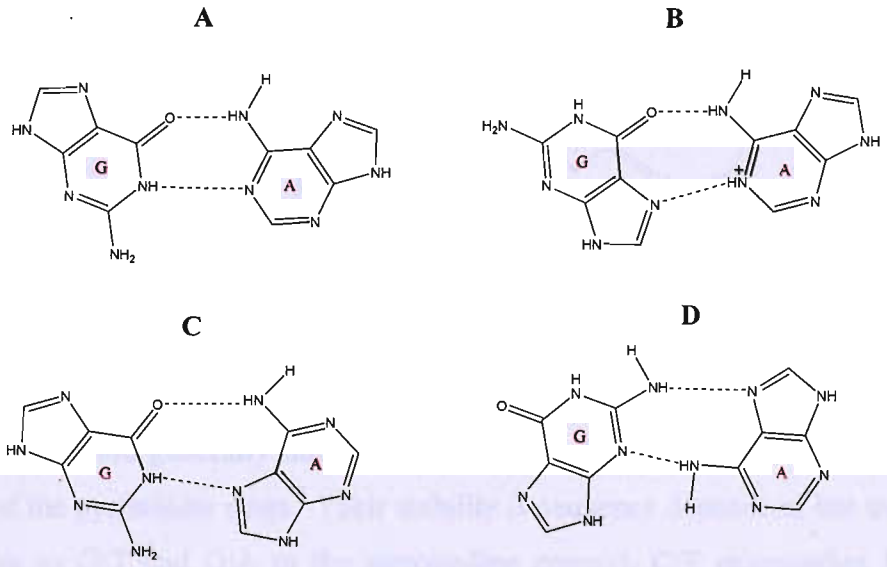


Figure 4.4 Structures of G:A mismatches at neutral pH (A, C and D) and at acidic pH (B). Adapted from ref. 1.

The G:A mismatch stability is strongly context dependent.[1] The structures usually adopted in solution are shown in Figures 4.4A and B, at neutral and acidic pH, respectively.[1, 14] The geometry adopted in Figure 4.4C has also been observed in crystallographic studies[14, 15] and the geometry in Figure 4.4D is often observed for tandem G:A mismatches.[16] According to Allawi *et al.*[1] the thermodynamics of single G:A mismatches do not exhibit a strong pH dependence.

C:T mismatches

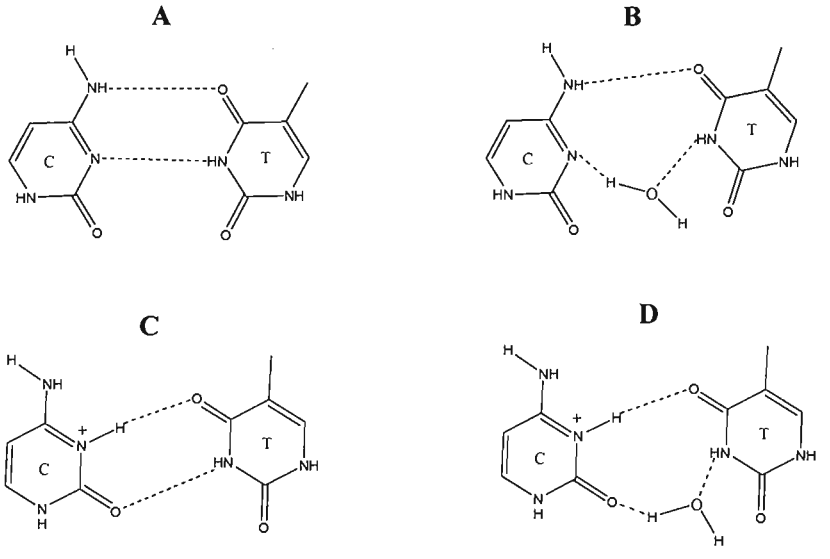


Figure 4.5 Structures of C:T mismatches at neutral pH (A, B) and at acidic pH (C, D). Adapted from ref. 7.

C:T mismatches are generally unstable mismatches. This is due to the weak stacking potential of the pyrimidine rings.[7] Their stability is sequence dependent, but they are not as sensitive as G:T and G:A to the surrounding context. C:T mismatches have been proposed to adopt at least four different structures depending on the sequence context and the pH of the medium (Figure 4.5). According to Allawi *et al.*[7] C:T mismatches both at neutral (Figure 4.5A and B) and acidic (Figure 4.5C and D) pH can pair with two hydrogen bonds one of which, due to the repulsion of the carbonyl groups of the cytosine and thymine, is possibly mediated via a water molecule. Also, according to Allawi, the thermodynamics of the C:T mismatch at lower pH (5) are slightly less stable than at neutral pH.

4.1.2 Stability trend for single base mismatches

As already said in previous sections the stability of the single base mismatches is context dependent. According to Peyret *et al.*[5] the stability trend averaged over all contexts, including the Watson-Crick base-pairs, is: G:C > A:T > G:G > G:T ≈ G:A > $A:C^+$ > T:T ≈ A:A ≈ $C:C^+$ > T:C ≥ A:C ≥ C:C.

The thermodynamic properties of nucleic acid duplex formation and dissociation in solution have been well established.[17] However, duplex formation, involving surface-immobilised DNA probes is less well understood due to the complex factors introduced by the solid phase, affecting the kinetics and thermodynamics of target capture. As mentioned before, care must be exerted when extrapolating solution thermodynamic predictions to the solid phase.

4.2 Aims of this chapter

Hybridisation on solid supports has not been studied as extensively as in solution.[17-19] In particular the relationship between fluorescence read-out, hybridisation strength and selectivity, and solution and washing conditions, although crucial to the performance of the array platform has not been studied in detail.

These issues were the driving force behind the experiments in this chapter. A comprehensive set of single base mismatches located centrally in 20-mer duplexes was designed to study the relationship between hybridisation selectivity, post-hybridisation washes and fluorescence read-out and how that correlates with solution melting temperatures. A multiplex assay involving the single base mismatched oligonucleotides was also explored.

4.3 Results and Discussion

4.3.1 T_m determination for a comprehensive set of single base mismatches

Figure 4.6 shows the oligonucleotide construct designed to study the stability of a comprehensive set of mismatches located centrally in the double helix. Four biotinylated probes differing in position X (ATD0958, ATD1548, ATD1549 and ATD1636) and four Cy5 labelled targets (ATD0960, ATD1201, ATD1202 and ATD1203) differing in position Y were synthesised. This allowed the study of all four possible perfectly matched combinations as well as the twelve possible mismatches. This experiment aimed to investigate the relationship between the stability of the duplexes in solution, fluorescence read-out of the beads (immobilised duplexes) and post-hybridisation washing conditions.

Figure 4.6 DNA sequences and schematics for the constructed X:Y base pairs. Position X refers to the varying base on the probe while Y refers to the varying base on the target.

The solution T_m of all the DNA duplexes was determined in a range of different conditions representing different stages of washing the beads (5xSSPE, 0.1 % SDS and 50 mM, 1 M, 2 M urea). Table 4.1 shows the thermal stability variation of each X:Y base pair with increasing stringency of the solvent used. Figure 4.7A and B show a graphical representation of the results shown on table 4.1 for the base pairs C:Y and A:Y.

Base pair X:Y	T_m (5xSSPE, 0.1 % SDS)	T_m (50 mM urea)	T_m (1 M urea)	T_m (2 M urea)
C:G	67.8±0.1	43.6±0.6	39.4±0.4	38.1±0.5
C:T	57.0±0.1	31.5±0.1	28.3±0.4	25.3±0.4
C:A	57.4±0.1	33.1±0.2	29.6±0.4	28.9±0.4
C:C	54.9±0.2	33.0±0.2	26.5±0.7	25.6±0.9
G:C	66.3±0.2	42.1±1.0	40.1±0.3	38.1±0.4
G:A	60.9±0.1	37.2±0.8	34.2±0.5	31.9±0.5
G:T	60.1±0.1	36.3±1.0	33.7±0.5	31.7±0.4
G:G	62.7±0.1	38.4±0.8	35.7±0.3	33.1±0.3
T:A	64.9±0.1	39.6±1.2	37.4±0.5	35.9±0.4
T:T	57.3±0.2	32.7±1.1	30.5±0.3	28.5±0.3
T:G	60.2±0.1	35.8±1.0	33.7±0.4	31.6±0.3
T:C	56.4±0.1	32.2±0.8	28.6±0.6	26.7±0.4
A:T	64.8±0.2	39.4±1.3	37.5±0.5	36.0±0.5
A:G	62.5±0.2	36.7±1.1	34.3±0.6	32.7±0.5
A:A	58.7±0.2	35.3±1.0	31.3±0.6	28.8±0.6
A:C	56.7±0.1	32.7±0.9	29.7±0.4	27.5±0.5

Table 4.1 T_m (± SE) determinations for the X:Y base pairs. Watson-Crick base pairs are highlighted in red.

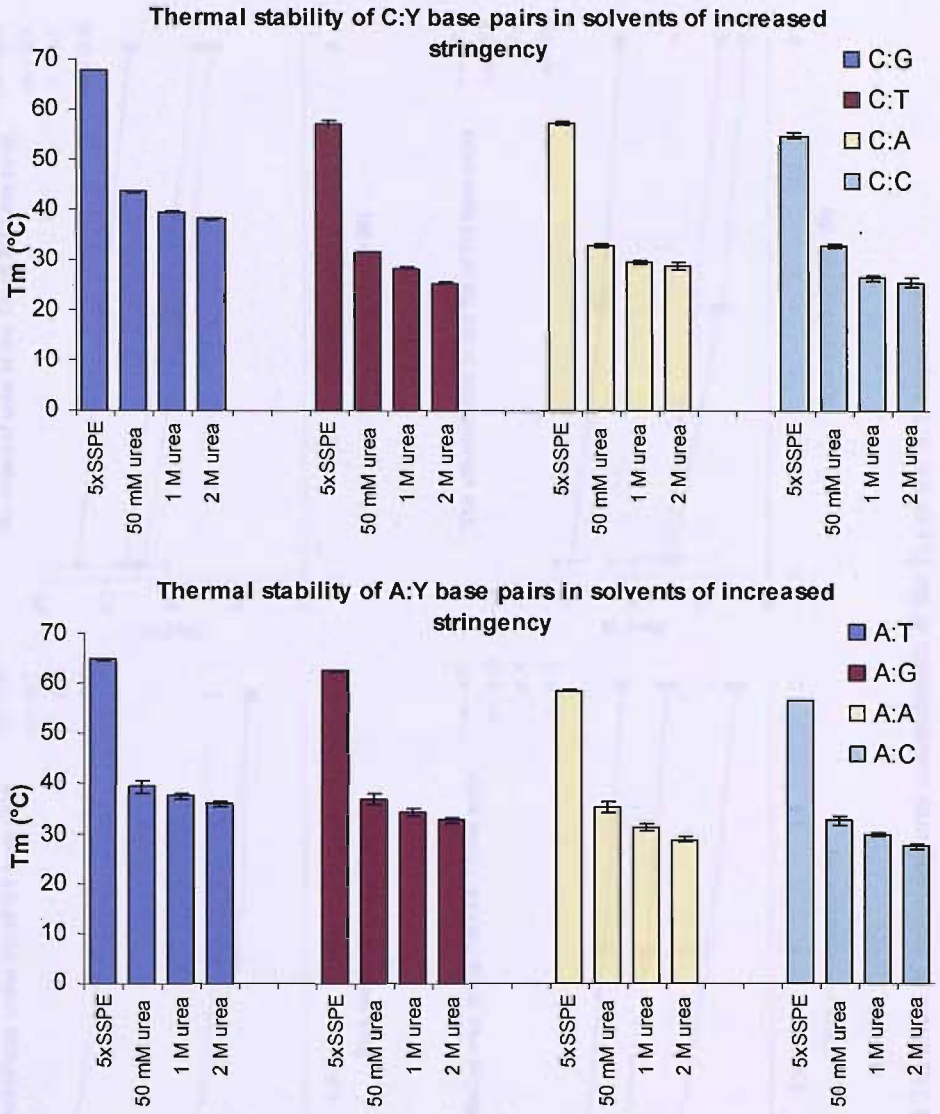


Figure 4.7 Graphical representation of T_ms determined for C:Y and A:Y base pairs.

For oligonucleotides approximately 15 to 20 bases in length, the T_m for the hybridisation of a perfectly matched duplex compared to one with a single base mismatch can differ by several degrees.[20] The results shown in Table 4.1 are consistent with this evidence, as the matched sequences exhibited higher T_m than their respective mismatches in all solvents studied. All T_ms decreased steadily, as expected, with increasing solvent stringency[21] but remained above room temperature. The increase in urea concentration led to an approximately linear decrease in T_m in agreement with the findings of Hutton[22] (Figure 4.8).

A closer inspection into the T_ms obtained across solvents, reveals that the mismatches containing guanine displayed higher T_ms than the rest of the mismatched base pairs, as anticipated.[10, 12] The stability trend published by Peyret *et al.*[5] G:C > A:T > G:G > G:T ≈ G:A > T:T ≈ A:A > T:C ≥ A:C ≥ C:C averaged all over sequence contexts, correlates very well with the results presented on table 4.1, in all solvents. The only difference is that the stability relationship between T:C (C:T) and A:C (C:A) inverted in urea solvents.

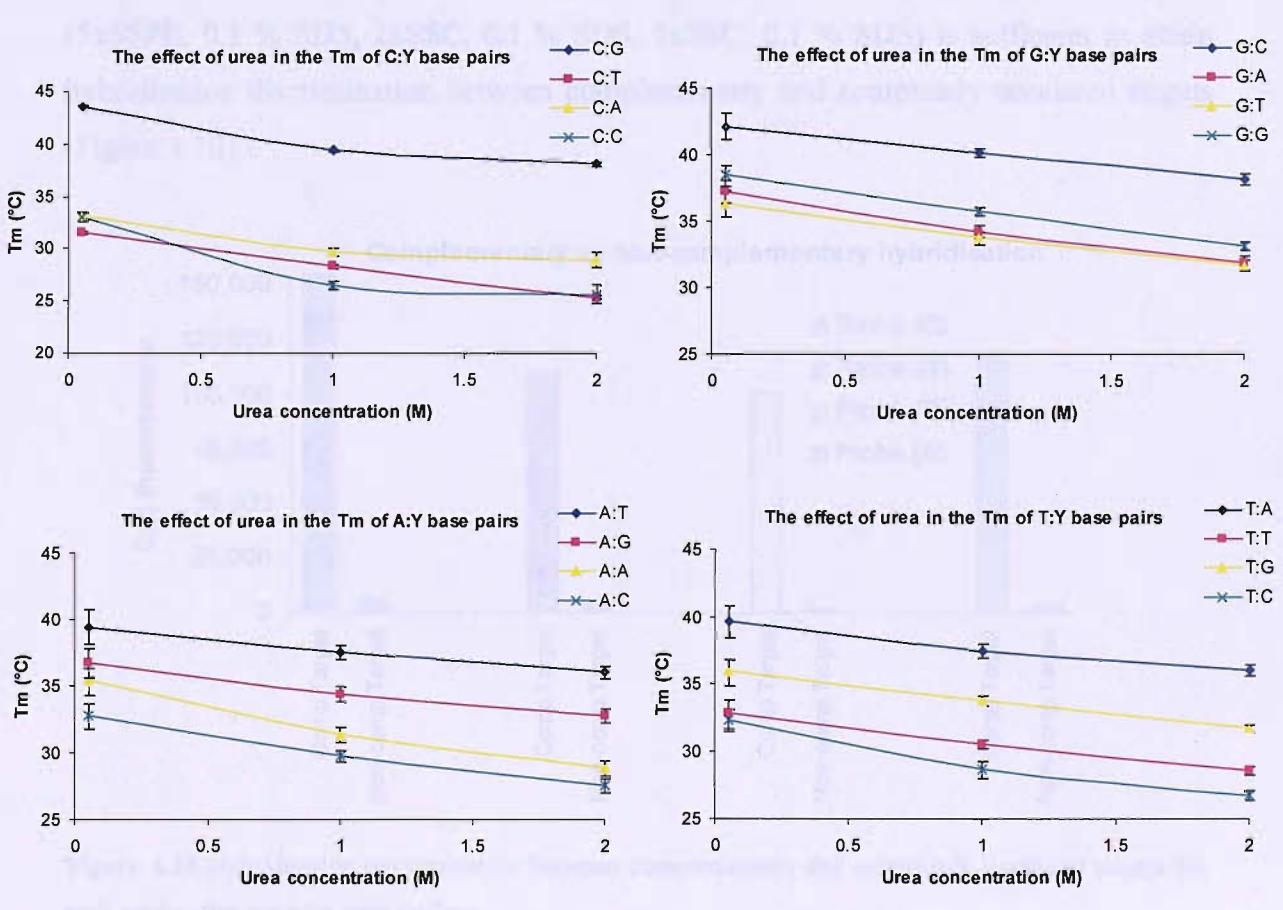


Figure 4.8 The effect of increasing urea concentration in the T_ms of the X:Y sequences.

4.3.2 Fluorescence read-out of the immobilised single base mismatches

The next step of the experiment was to examine the effect of the different X:Y mismatches on the selectivity of hybridisation, under the post-hybridisation washing conditions. The biotinylated 20-mer probes were immobilised onto avidin coated GMA beads and incubated overnight at room temperature with the Cy5 labelled targets (Figure 4.9). The beads were then washed with solutions of increased stringency, to remove non-specifically bound targets and analysed by FACS at each stage (Table 4.2). The level of non-complementary hybridisation of the probes to a completely unrelated Cy5 labelled 30-mer target (4goligo0016) was first assessed (Figure 4.10).

Figure 4.9 Schematics of the immobilised DNA duplexes.

As already seen in chapter 3, washing the beads with buffers of increased stringency (5xSSPE, 0.1 % SDS, 2xSSC, 0.1 % SDS, 1xSSC, 0.1 % SDS) is sufficient to attain hybridisation discrimination between complementary and completely unrelated targets (Figure 4.10).

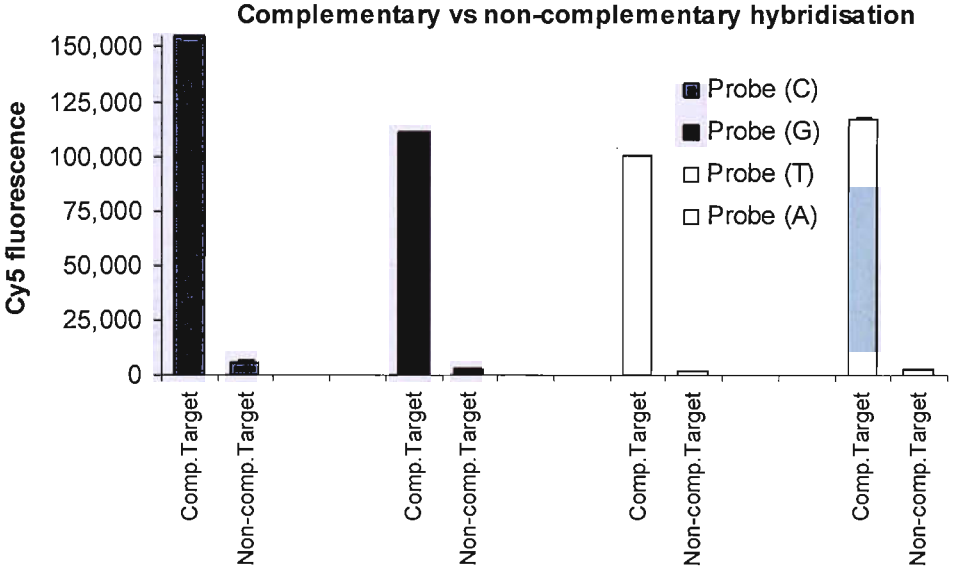


Figure 4.10 Hybridisation discrimination between complementary and completely unrelated targets for each probe after washing with buffers.

Base pair X:Y	5xSSPE, 0.1% SDS	2xSSC, 0.1% SDS	1xSSC, 0.1% SDS	50 mM urea	1 M urea	2 M urea
C:G	216,438±419.8	176,345±473.2	154,845±500.0	128,602±477.3	98,820±434.3	53,898±350.1
C:T	180,836±434.9	127,604±396.9	107,936±373.0	81,344±261.9	31,614±134.6	4,857±74.1
C:A	197,425±414.7	158,110±463.2	130,587±439.4	79,048±419.8	47,225±329.5	9,876±103.0
C:C	173,087±376.1	151,671±462.5	121,056±109.6	57,381±385.0	28,985±251.8	9,745±153.9
G:C	173,283±326.3	127,138±271.8	110,348±261.3	96,889±219.7	77,866±226.9	44,211±200.9
G:A	124,813±255.4	120,977±298.2	107,872±269.5	82,525±192.6	54,307±157.5	8,983±84.0
G:T	165,476±384.6	122,402±301.4	116,607±270.1	92,420±225.6	51,500±158.6	6,043±57.7
G:G	145,105±345.5	111,437±281.6	97,983±257.4	71,963±210.7	51,666±197.8	14,836±91.3
T:A	170,396±362.9	119,304±249.2	100,394±235.5	87,083±223.4	70,488±226.3	55,438±196.4
T:T	140,054±329.2	111,965±266.4	97,604±249.6	77,038±207.4	51,410±167.3	6,257±56.4
T:G	138,690±346.0	116,617±256.9	108,161±231.8	85,821±212.4	48,897±141.2	8,070±65.1
T:C	167,745±348.6	104,346±271.8	89,025±234.1	67,924±226.0	47,739±165.8	6,418±65.7
A:T	185,225±396.8	136,798±403.2	117,304±384.1	87,588±289.7	67,837±276.2	52,407±236.0
A:G	146,862±533.9	139,143±514.6	131,546±528.6	105,129±479.4	84,238±481.6	19,645±233.7
A:A	143,485±507.8	135,842±493.4	128,229±504.4	84,485±472.9	61,898±449.9	8,291±155.3
A:C	137,727±401.3	121,255±397.2	118,117±404.3	33,906±344.5	48,285±302.5	8,077±115.1

Table 4.2 Fluorescence (± SE) read-out of the beads for the X:Y base pairs with increasing washing stringency. Watson-Crick base pairs are highlighted in red. The mismatches highlighted in blue, after washing with 1xSSC, 0.1 % SDS, show higher fluorescence than the corresponding complementary base pairs.

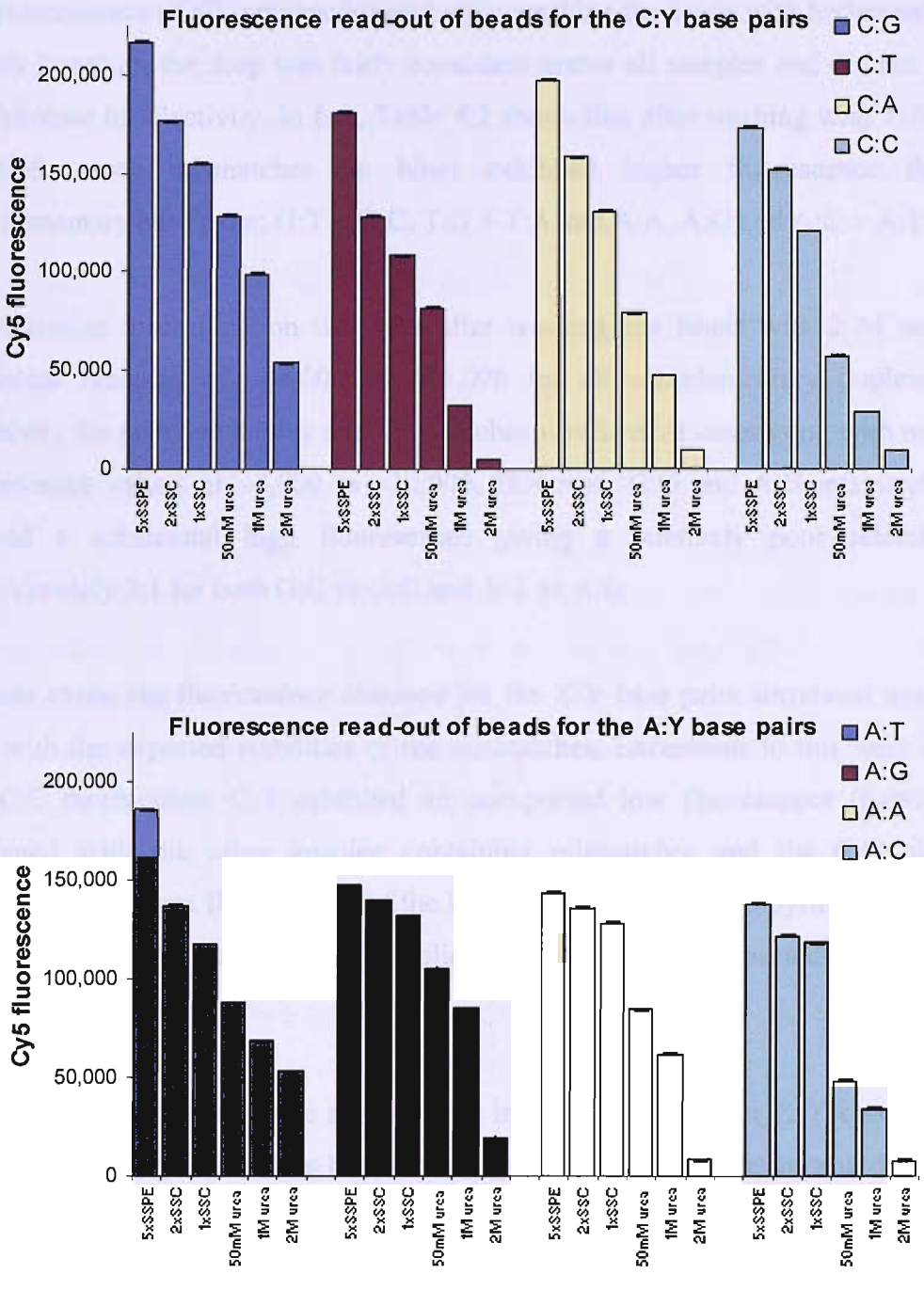


Figure 4.11 Graphical representation of the fluorescence (± SE) read-out of C:Y and A:Y base pairs with increasing washing stringency.

The results shown on Table 4.2 and the fluorescence profiles in Figure 4.11, show that after washing with hybridisation buffer (5xSSPE, 0.1 % SDS) the fluorescence of all samples was very similar for each probe. In this context, the highest fluorescence was exhibited for the C:G duplex and the lowest was observed for T:G and A:C mismatches.

The fluorescence of all samples dropped upon washing the beads with higher stringency buffers however, the drop was fairly consistent across all samples and did not provide any increase in selectivity. In fact, Table 4.2 shows that after washing with 1xSSC, 0.1 % SDS, some mismatches (in blue) exhibited higher fluorescence than the complementary base pairs; G:T > G:C, T:G > T:A and A:A, A:G and A:C > A:T.

Hybridisation discrimination was best after washing the beads with 2 M urea with consistent readings of ~44,000 to ~55,000 for all complementary duplexes. The selectivity for matched duplex over mismatches was in most cases good with mismatch fluorescence values of ~4,000 to ~10,000. However, G:G and A:G mismatches still showed a substantial high fluorescence giving a relatively poor selectivity of approximately 3:1 for both G:C vs G:G and A:T vs A:G.

In most cases, the fluorescence obtained for the X:Y base pairs correlated moderately well with the expected stabilities of the mismatches. Exceptions to this were the G:T and C:C mismatches. G:T exhibited an unexpected low fluorescence (6,043) when compared with the other guanine containing mismatches and the C:C mismatch exhibited the highest fluorescence of the base pairs formed by two pyrimidines (9,745). However, the C:C stability on the solid phase was not unexpected[9] and was also observed in chapter 3.

The results also showed some asymmetries in signal intensity for X:Y base pairs that contain the same pair of bases but differ only in the sense that the mismatched base is either on the probe or target of the duplex. Obvious examples of this were, C:G and G:C, A:G and G:A, G:T and T:G and T:C and C:T. This was consistent across experiments and according to Pozhitkov *et al.*[17] who also observed similar signal asymmetries on a planar microarray, cannot be currently explained.

Varying the temperature of the hybridisation (55 °C and room temperature) and hybridisation time from fifteen minutes to overnight had no effect on the selectivity achieved. However, the order of the washing steps was important. If the intermediate washing steps were skipped, washing the beads in 2 M urea would lead to rapid loss of fluorescence signal for both matched and mismatched duplexes without any selectivity achieved (data not shown).

4.3.3 Solution T_m is a poor predictor of fluorescence read-out

Previous sections have shown that the fluorescence read-out of the beads and the solution T_m of the duplexes did not follow the same stability trend for the X:Y base pairs. This was not unexpected as Pozhitkov *et al.* in a recent publication observed that the order of stabilities of single base mismatches on a microarray assay was different from that exhibited in solution.[17] The authors found that regardless of the position of the mismatch on 20-mer duplexes, pyrimidine-pyrimidine mismatches were in general more stable than purine-purine mismatches. However, this is not in agreement with the fluorescence read-out results shown on Table 4.2.

Figure 4.12 shows the relationship between the solution T_ms and the fluorescence read-out observed. The correlation between solution T_ms and the fluorescence read-out in hybridisation buffer (5xSSPE, 0.1 % SDS) is very poor (Figure 4.12A). The correlation of fluorescence after washing with 2 M urea with T_m in either buffer or urea solutions is apparently improved (Figures 4.12B and C).

Figures 4.12B and C show graphs with two distinct regions; one at higher T_m and higher fluorescence, indicating the perfectly matched base pairs and another region located at lower T_m and lower fluorescence indicating the X:Y mismatches. However, this can be largely credited to the ability of the improved discrimination between all matched and all mismatched duplexes after washing with 2 M urea. It was not a decrease in thermodynamic stability that was giving the observed hybridisation discrimination. If it were, the fluorescence read-out would have shown an X:Y stability trend similar to that exhibited by the solution T_ms. Overall, T_m was considered to be a fairly poor predictor of fluorescence hybridisation signal. This is in agreement with the findings of Pozhitkov *et al.*[17] The authors found that thermodynamic parameters failed to predict signal intensity values of duplexes with RNAs on an oligonucleotide (20 to 25-mers) DNA microarray.

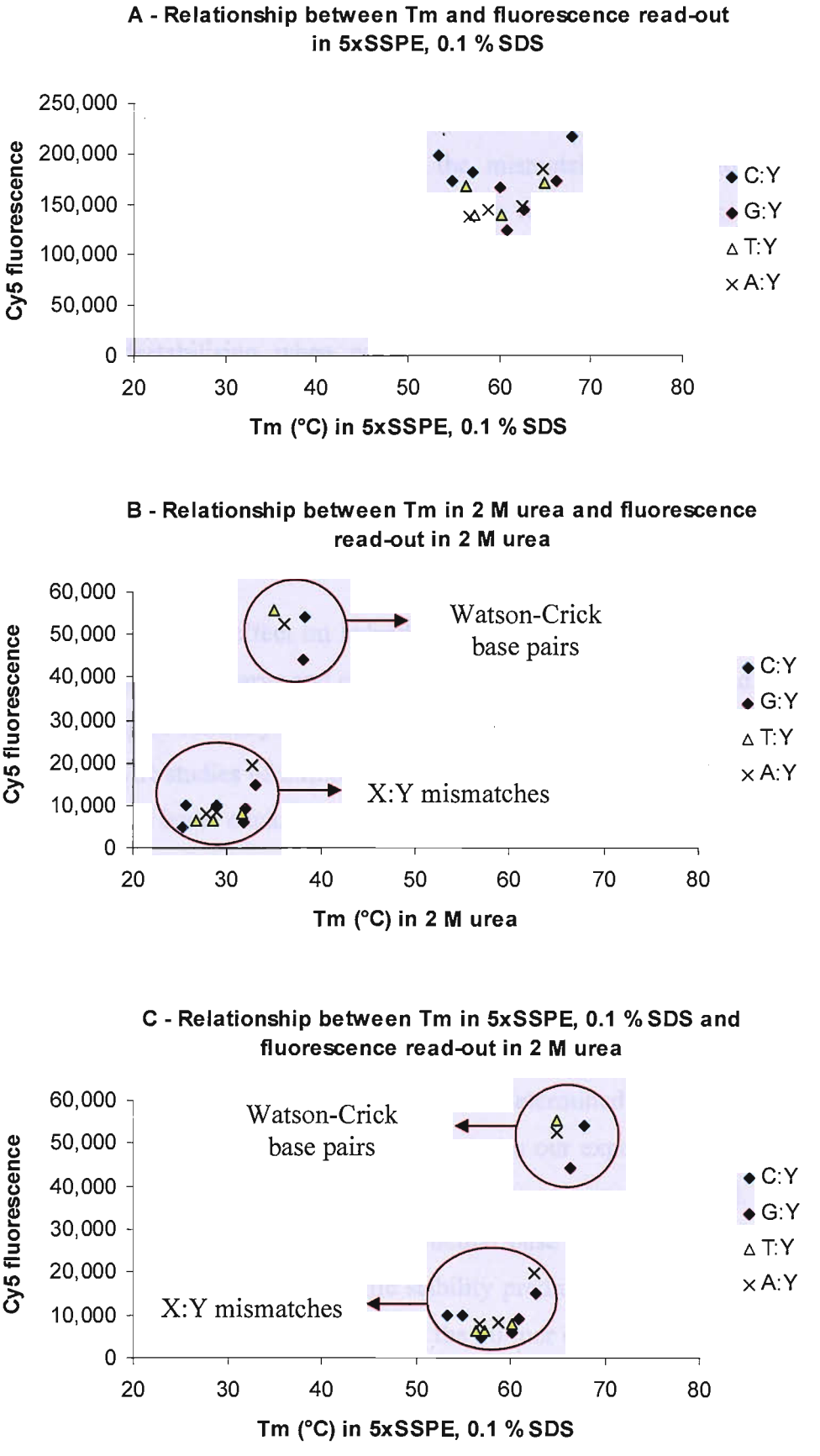


Figure 4.12 T_m correlations with the fluorescence of the beads in 5xSSPE, 0.1% SDS (A) and 2 M urea (B) respectively. (C) Correlation between the T_m determined in 5xSSPE, 0.1% SDS and the fluorescence read-out of the beads in 2 M urea.

This posed the question as to whether a kinetic effect rather than a thermodynamic effect might be responsible for the stability results observed on solid phase. The results show that the duplexes strongly dissociated in urea solutions. This dissociation was higher for the mismatches. This suggested the mismatches had faster kinetics of dissociation than the Watson-Crick matches under the washing conditions. According to Bhattacharya *et al.*,[3] mismatches produce local destabilisation in the helical duplex accompanied by a local structure distortion, disrupting the π stacked array, rendering them kinetically destabilising when compared to the Watson-Crick matches. The mismatched base pairs are thus more accessible to solvent interactions than the corresponding Watson-Crick matches. This would explain why the interaction of urea with the mismatched double helixes led to faster kinetics of dissociation for those duplexes.

The possibility of a kinetic effect on hybridisation discrimination was reinforced by a study carried out by Bhattacharya and co-workers.[3] The authors investigated the kinetics of base pair opening of centrally located homologous mismatches (AXT:TYA) in 9-mer duplexes by ^{1}H NMR studies of imino proton exchange rates. The imino protons of the aromatic heterocyclic bases exchange protons with the solvent when the hydrogen bond in the base pair is disrupted. This solvent accessibility which causes the proton exchange is faster for mismatches and consequently they have shorter lifetimes than matched base pairs. The kinetics of base pair opening trend found was GG > AA > CC > TT, with G:G displaying the longest base pair lifetime. This trend correlates very well with the fluorescence read-out for the homologous mismatches after washing with 2 M urea (GG > CC > AA > TT). The T_ms stability trend determined in 5 mM Na_2HPO_4, 15 mM NaCl, pH 7.0 by the authors is in agreement with our experimental results, GG > TT ≈ AA > CC and contrasts with the kinetics of base pair opening trend. This shows that the kinetics of base pair opening of a particular base pair under a given sequence context may not reflect the thermodynamic stability predicted by the T_m. According to these findings, a similar ^{1}H NMR study for all the 20-mer duplexes used in this chapter may be a good predictor of the fluorescence read-out observed.

4.3.4 Another poor correlation between T_m and fluorescence read-out

The fact that T_m is a poor predictor of fluorescence hybridisation signal was further reinforced by a study carried out on the oligonucleotide constructs shown in Figure 4.13, containing a centrally located double mismatch. The duplexes refer to a HIV wild type protease and V82A, a drug resistant mutant and their complementary Cy5 labelled targets.

Figure 4.13 Schematics of a HIV wild type protease and V82A mutation sequences.

Table 4.3 shows the solution T_ms determined in 5xSSPE, 0.1 % SDS and 2 M urea for the complementary and mismatched hybridisations of probes HIV wild type and V82A. As would be expected for both probes, the double mismatch leads to a significant decrease in T_m. This is verified in both solvents.

Probe	Hybridisation	Duplex sequence	T_m (in 5xSSPE, 0.1 % SDS)	T_m (in 2 M urea)
HIV wt	Complementary	5'…GTCA…3' 3'…CAGT…5'	66.9±0.4	38.2±0.6
HIV wt	Mismatched	5'…GTCA…3' 3'…CGAT…5'	54.7±0.2	28.7±1.7
V82A	Complementary	5' …GCTA…3' 3'…CGAT…5'	67.8±0.1	38.1±0.5
V82A	Mismatched	5' …GCTA…3' 3'…CAGT…5'	54.8±0.2	25.3±0.2

Table 4.3 T_ms (± SE) for the complementary and mismatched (in red) hybridisation of HIV wt and V82A.

However, for probe HIV wild type, the mismatched hybridisation gave as higher fluorescence signal as the complementary hybridisation even after washing the beads with 2 M urea (Figure 4.14). The stringency was increased up to 8 M urea (not shown), but selectivity was not achieved. Therefore this mismatch could not be distinguished. This was surprising and shows a complex interaction between the sequence mutation and the washings. For probe V82A however, hybridisation discrimination was successfully achieved.

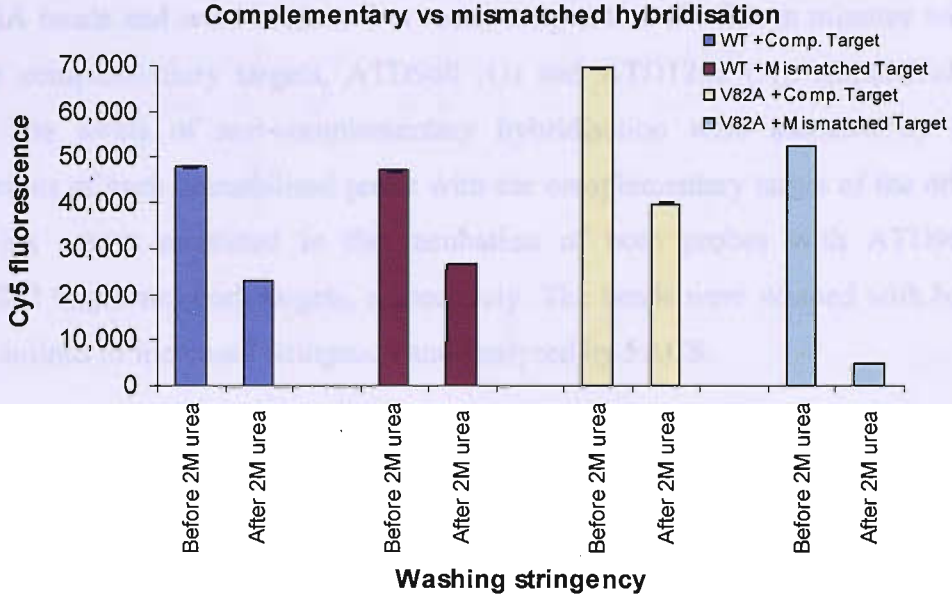


Figure 4.14 Fluorescence (± SE) read-out before and after washing the beads with 2 M urea, for the complementary and mismatched hybridisations of probes HIV wild type and V82A.

4.3.5 Multiplex assay: assessing the cross-reactivity of two immobilised probes differing in one base and two targets in solution

Two of the probes mentioned in section 4.3.1, differing in one base and their targets were chosen to investigate a multiplexed assay and compare that with the individual hybridisation results. A fluorescent probe, fluorescent targets and FRET used as a marker for correct hybridisation, were utilised as explained in Chapter 3.

Both 20-mer 5' biotinylated probes, ATD1200 (C) and ATD1549 (T), were immobilised on GMA beads and were incubated at room temperature for fifteen minutes with 5' Cy5 20-mer complementary targets, ATD960 (G) and ATD1202 (A), respectively (Figure 4.15). The levels of non-complementary hybridisation were assessed by individual incubations of each immobilised probe with the complementary target of the other probe. Multiplex assays consisted in the incubation of both probes with ATD960 target, ATD1202 target and both targets, respectively. The beads were washed with buffers and urea solutions of increased stringency and analysed by FACS.

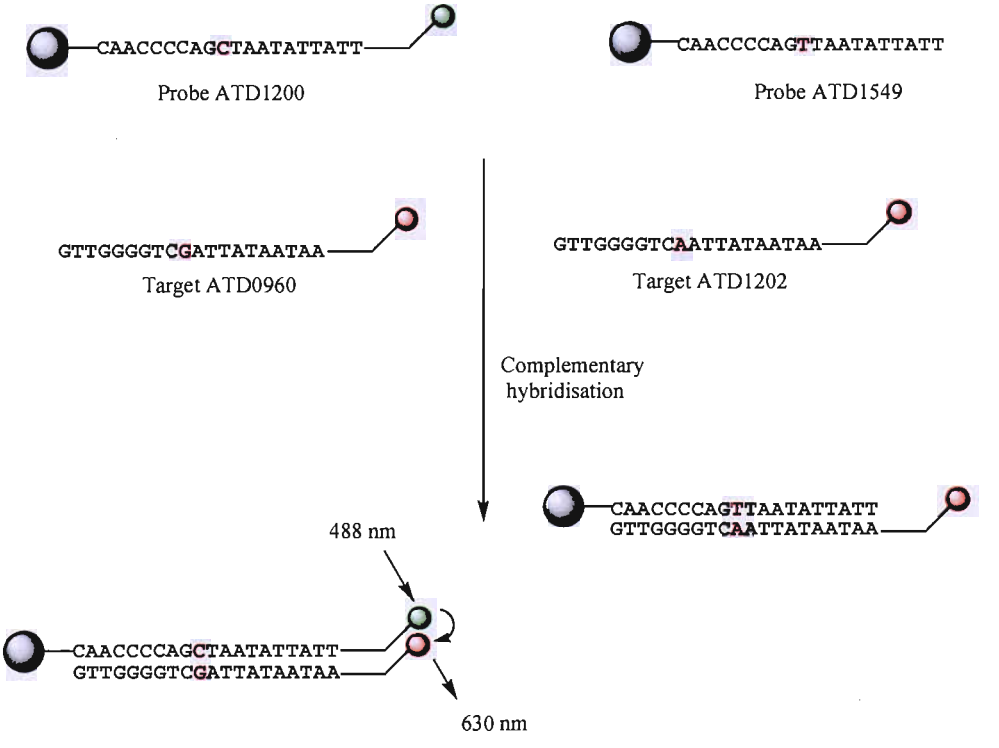


Figure 4.15 Schematics of the complementarity between probes and targets. The complementary hybridisation of probe ATD1200 to target ATD0960 generates FRET, due to the close proximity of the fluorophores.

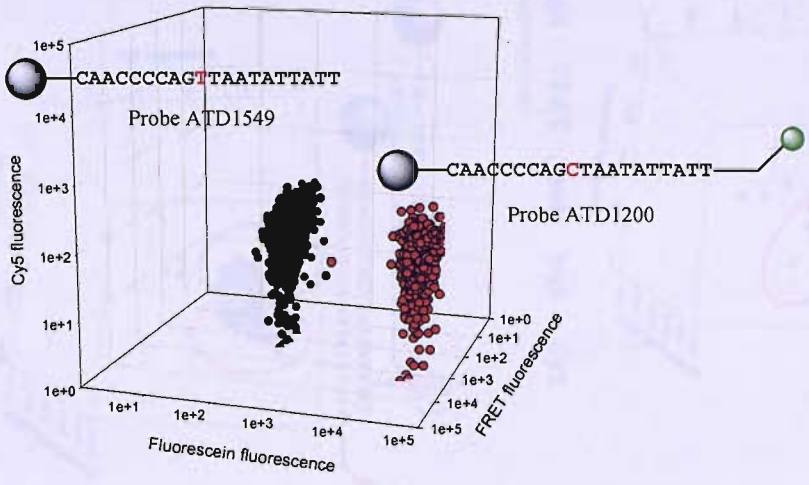
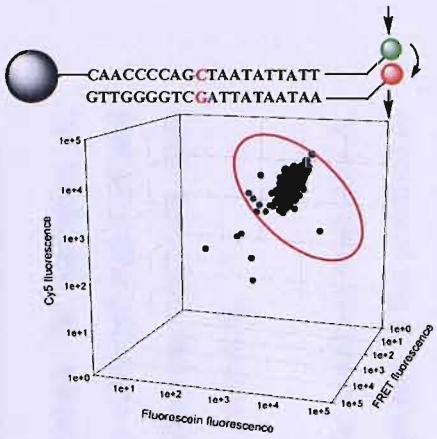


Figure 4.16 Three-dimensional representation of Cy5, fluorescein and FRET fluorescence exhibited by the populations of probes ATD1200 and ATD1549 (The fluorescence read-out of each probe was obtained individually).

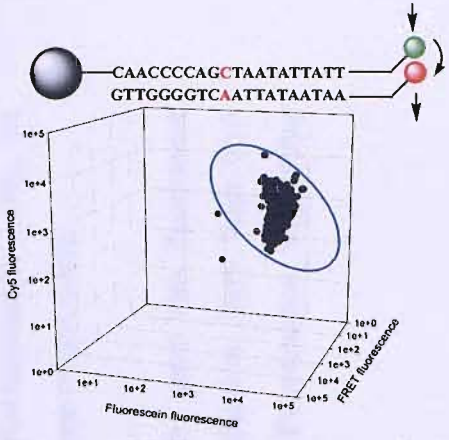
Figure 4.16 shows that probe ATD1200 (C) exhibited very high fluorescein fluorescence and minimal Cy5 and FRET background, as expected. On the other hand, non-labelled probe ATD1549 (T) exhibited minimal fluorescence in all three channels. Figure 4.17 shows the results obtained for the complementary and mismatched hybridisation reactions and multiplex assays.

The hybridisation of probe ATD1200 (C) to 5' Cy5 complementary target, ATD0960 (G), resulted in a population with high Cy5 and FRET fluorescence and low fluorescein fluorescence, as anticipated. The incubation of probe ATD1200 (C) to mismatched target ATD1202 (A) generated a population with high fluorescein fluorescence and low Cy5 and FRET fluorescence suggesting minimal mismatched hybridisation.

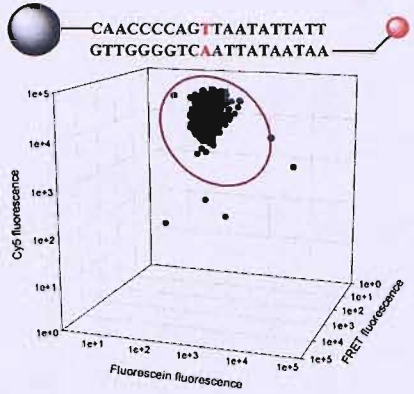
The hybridisation of probe ATD1549 (T) to 5' Cy5 complementary target, ATD1202 (A), generated a population with very high Cy5 fluorescence, as expected. The incubation with mismatched target, ATD0960 (G), exhibited a population high in Cy5 fluorescence, suggesting T:G mismatched hybridisation.



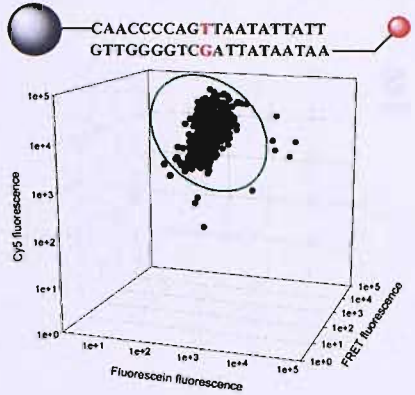
High FRET and high Cy5 fluorescence



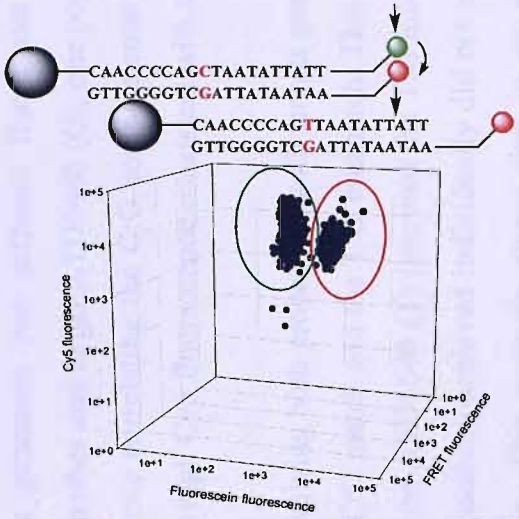
Low FRET and low Cy5 fluorescence, indicating virtually no mismatched hybridisation



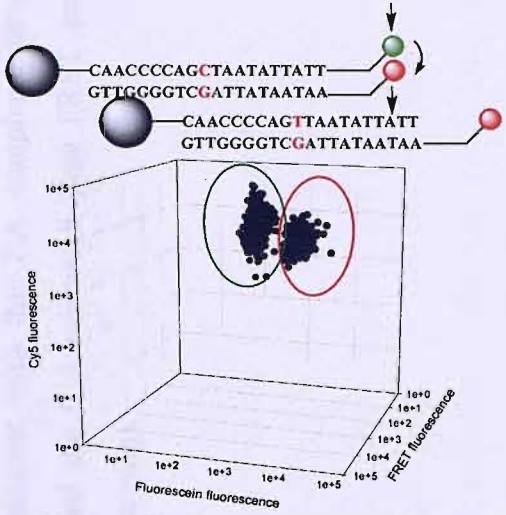
Very high Cy5 fluorescence



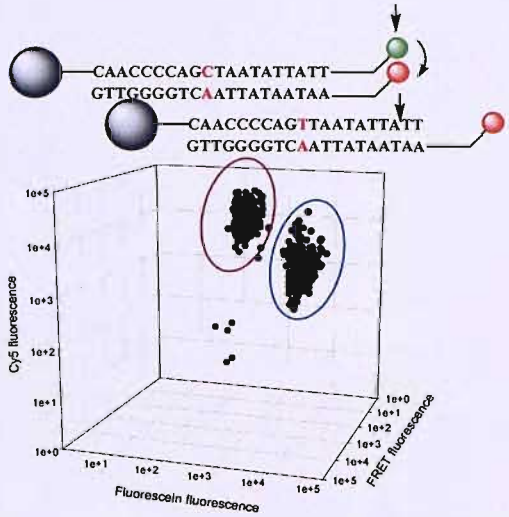
High Cy5 fluorescence, indicating T:G mismatched hybridisation



Multiplex A - both probes and target ATD0960. Green population exhibits fluorescence consistent with T:G mismatching



Multiplex B - both probes and both targets. Green population exhibits fluorescence consistent with T:G mismatching



Multiplex C - both probes and target ATD1202. Blue population exhibits low FRET and low Cy5 fluorescence

Figure 4.17 Fluorescence of the beads for the complementary and mismatched individual hybridisation reactions and for the multiplex assays. Three-dimensional plots showing Cy5, fluorescein and FRET fluorescence.

All multiplexes generated two different fluorescent populations. In multiplex A, involving both probes and target ATD0960 (G), one population exhibited high Cy5 and FRET fluorescence correlating the C:G complementary hybridisation, while the other population gave high Cy5 fluorescence consistent with the T:G mismatch.

Multiplex B, involving both probes and both targets gave similar results to multiplex A, correlating the C:G match and the T:G mismatch. This suggests that in the presence of both targets, probe ATD1549 (T) favoured the T:G mismatch over the T:A match. The hybridisation selectivity achieved individually did not correlate the selectivity observed in the multiplex. This example shows that in an allelic discrimination assay, an interference caused by a stable mismatch could mask the presence of one allele.

Multiplex C, involving both probes and target ATD1202 (A) gave two fluorescent populations consistent with the T:A complementary hybridisation (very high Cy5 fluorescence) and the C:A mismatch (low FRET and Cy5 fluorescence, suggesting minimal mismatching).

4.4 Conclusions

This chapter aimed to understand the relationship between solution T_ms and hybridisation selectivity and fluorescence read-out and the role of post-hybridisation washes in the context of suspension arrays. For that a set of four 20-mer probes and four Cy5 labelled 20-mer targets were synthesised. These probes and targets possessed a varying base located in the middle of the sequence. This allowed the study of a comprehensive set of X:Y mismatches.

The solution T_ms of oligonucleotide duplexes were determined experimentally in a range of solvents of increased stringency representing different stages of the post-hybridisation washes. As anticipated, the increase in stringency decreased dramatically the T_ms of the duplexes. All matched duplexes exhibited higher T_ms than the mismatches and all the T_ms determined remained above room temperature. The stability trend obtained for all X:Y base pairs was consistent in all solvents and was in very good agreement with the literature prediction.

The fluorescence read-out was also determined experimentally for all immobilised duplexes. However, the stability trend observed on the solid phase for the X:Y base pairs was not the same observed by the solution melting temperatures. The results obtained at each step of the post-hybridisation washes, reinforced the need to use buffers and urea solutions of increased stringency to attain selectivity of hybridisation for closely related sequences. The order of the washing steps was important. Selectivity was best after washing the beads with 2 M urea. However, some mismatches, G:G and A:G still exhibited a high fluorescence intensity (3:1) when compared to the respective matches. Also, asymmetries in the fluorescence read-out were observed for some X:Y combinations containing the same base pairs; C:G / G:C, A:G / G:A, G:T / T:G and T:C / C:T.

The correlation between solution T_ms and the fluorescence of the beads was seen to be very poor in hybridisation buffer. The correlation between the T_ms and the fluorescence of the beads after washing in 2 M urea showed a distinction between all matched and all mismatched duplexes. However, it is believed that the distinction achieved between all matches and all mismatches was due to the higher rate of dissociation of the mismatches

under the washing conditions employed, rather than discrimination based in thermodynamic stability. Overall, T_m was considered to be a poor predictor of fluorescence hybridisation signal. This was most distinctly seen in the case of one anomalous duplex containing a central double mismatch. The presence of the double mismatch led to a significant decrease in the experimentally determined T_m when compared to the corresponding fully matched duplex. However, the double mismatch generated a similar fluorescence read-out as the matched duplex and could not be distinguished.

The possibility of a kinetic effect responsible for the hybridisation discrimination observed for the X:Y base pairs was reinforced by a study carried out by Bhattacharya and co-workers on homologous mismatches.[3] The authors showed that the thermodynamic stability predicted by the T_m for a given base pair may not be reflected on the kinetics of base pair opening. The stability trend exhibited by the solution T_ms determined by the authors for the homologous mismatches (in agreement with ours) did not correlate the kinetics of base pair opening of the mismatches (in very good agreement with fluorescence read-out) as determined by ^{1}H NMR.

One of the multiplex assays examined indicated that a probe (T) in the presence of two targets (containing G and A, respectively) favoured the T:G mismatch over the T:A match. This result did not correlate the selectivity of hybridisation observed in the individual reactions. This raises the concern that in a SNP discrimination assay, the formation of a stable mismatch might mask the presence of one allele.

The results presented in this chapter show that the stability exhibited by a mismatch in the context of an array may differ from that predicted by solution-based hybridisations. The notion that short oligonucleotides with a mismatch should hybridise less efficiently than a perfect match in the context of arrays[17, 23] is not always applicable. The type of mismatch, location and sequence context on hybridisation signal are far from understood.[17] This makes difficult to predict the hybridisation performance of oligonucleotide probes and has serious implications on probe design.

4.5 Experimental methods

4.5.1 Reagents

Amino-modified GMA beads (5 - 6 μm - 10% (W/V)) were obtained from Bangs Labs. Oligonucleotide sequences were synthesised by ATDbio and Sigma-Aldrich (Tables 4.4 and 4.5). Avidin DN (1.0 mg/ml) was from Vector Laboratories. EDC (1-ethyl-3-(3-dimethylaminopropyl) carbodiimide), sodium chloride, sodium citrate, sodium dihydrogen phosphate, EDTA, SDS (sodium dodecyl sulphate) and urea were obtained from Acros and imidazole came from Aldrich.

DNA Probes	Type of modification	Sequence 5' – 3'	Length (nt)
ATD0957	5' Biotin	CAACCCCAGTCAATATTATT	20
ATD0958	5' Biotin	CAACCCCAGCTAATATTATT	20
ATD1548	5' Biotin	CAACCCCAGGTAATATTATT	20
ATD1549	5' Biotin	CAACCCCAGTTAATATTATT	20
ATD1636	5' Biotin	CAACCCCAGATAATATTATT	20
ATD1200	5' Biotin 3' Fluorescein	CAACCCCAGCTAATATTATT	20

Table 4.4 DNA probe sequences employed.

DNA Targets	Type of modification	Sequence 5' – 3'	Length (nt)
ATD0960	5' Cy5	AATAATATTAGCTGGGGTTG	20
ATD1201	5' Cy5	AATAATATTATCTGGGGTTG	20
ATD1202	5' Cy5	AATAATATTAACTGGGGTTG	20
ATD1203	5' Cy5	AATAATATTACCTGGGGTTG	20
ATD0959	5' Cy5	AATAATATTGACTGGGGTTG	20
ATD1204	5' Cy5	AATAATATTACTGGGGTTG	19
4goligo0016	5' Cy5	CCATCGACTTGACTTACTCGAGTGCATCTC	30

Table 4.5 DNA target sequences employed.

4.5.2 T_m determination

Solution-phase melting temperatures of DNA duplexes were measured with a Cary Varian 400 SCAN UV-VIS spectrophotometer system with a Peltier thermostated twelve cell holder and motorised sample stage which allowed the simultaneous analysis of six samples and reference pairs using 1.5 ml quartz cuvettes. Each probe-target set contained an X:Y base pair combination as displayed in Figure 2.10 and Table 2.3 in 1ml of solvent. The DNA concentration of each strand was 1 μmol. The melting profile was then performed in 1°C per minute increments with constant monitoring at 260 nm for four gradient ramps. The T_m was taken as the average determined for each X:Y base pair combination for each gradient ramp by calculating the first derivative of the A260 profile.

4.5.3 FACS analysis

The experimental methodology for the FACS analysis of the samples of beads was performed as described in section 2.5.2.

4.5.4 Preparation of avidin functionalised solid supports

The experimental methodology for the avidin functionalisation of GMA beads is described in section 2.5.14.

4.5.5 Covalent attachment of biotinylated DNA to avidin coated GMA beads

The experimental methodology for the covalent attachment of biotinylated DNA probes to avidin coated GMA beads was as described in section 2.5.15.

4.5.6 DNA hybridisation on avidin coated GMA beads, multiplex assay and post-hybridisation stringency

The experimental methodology for DNA target hybridisation, multiplex assays and post-hybridisation washings was as described in sections 2.5.16, 3.5.5 and 3.5.6.

4.6 References

1. Allawi, H.T.; SantaLucia, J.Jr., Nearest neighbor thermodynamic parameters for internal G.A mismatches in DNA. *Biochemistry* **1998**, 37, 2170-9.
2. Allawi, H.T.; SantaLucia, J.Jr., Nearest-neighbor thermodynamics of internal A.C mismatches in DNA: sequence dependence and pH effects. *Biochemistry* **1998**, 37, 9435-44.
3. Bhattacharya, P.K.; Cha, J.; Barton, J.K., 1H NMR determination of base-pair lifetimes in oligonucleotides containing single base mismatches. *Nucleic Acids Res.* **2002**, 30, 4740-50.
4. Aboul-ela, F.; Koh, D.; Tinoco, I.Jr.; Martin, F.H., Base-base mismatches. Thermodynamics of double helix formation for dCA3XA3G + dCT3YT3G (X, Y = A,C,G,T). *Nucleic Acids Res.* **1985**, 13, 4811-24.
5. Peyret, N.; Seneviratne, P.A.; Allawi, H.T.; SantaLucia, J. Jr., Nearest-neighbor thermodynamics and NMR of DNA sequences with internal A.A, C.C, G.G, and T.T mismatches. *Biochemistry* **1999**, 12, 3468-77.
6. Isaacs, R.J.; Spielmann, H.P., Insight into GT mismatch recognition using molecular dynamics with time-averaged restraints derived from NMR spectroscopy. *J. Am. Chem. Soc.* **2004**, 126, 583-90.
7. Allawi, H.T.; SantaLucia, J.Jr., Thermodynamics of internal C.T mismatches in DNA. *Nucleic Acids Res.* **1998**, 26, 2694-701.
8. Tikhomirova, A.; Beletskaya, I.; Chalikian, T.V., Stability of DNA duplexes containing GG, CC, AA, and TT mismatches. *Biochemistry* **2006**, 45, 10563-71.
9. Wu, C.; Carta, R.; Zhang, L., Sequence dependence of cross-hybridization on short oligo microarrays. *Nucleic Acids Res.* **2005**, 33, e84.
10. Ke, S.H.; Wartell, R.M., Influence of nearest neighbor sequence on the stability of base pair mismatches in long DNA; determination by temperature-gradient gel electrophoresis. *Nucleic Acids Res.* **1993**, 22, 5137-43.
11. Faibis, V.; Cognet, J.A.; Boulard, Y.; Sowers, L.C.; Fazakerley, G.V., Solution structure of two mismatches G.G and I.I in the K-ras gene context by nuclear magnetic resonance and molecular dynamics. *Biochemistry* **1996**, 35, 14452-64.
12. Allawi, H.T.; SantaLucia, J.Jr., NMR solution structure of a DNA dodecamer containing single G.T mismatches. *Nucleic Acids Res.* **1998**, 21, 4925-34.
13. Allawi, H.T.; SantaLucia, J.Jr., Thermodynamics and NMR of internal G.T mismatches in DNA. *Biochemistry* **1997**, 36, 10581-94.
14. Webster, G.D.; Sanderson, M.R.; Skelly, J.V.; Neidle, S.; Swann, P.F.; Li, B.F.; Tickle, I.J., Crystal structure and sequence-dependent conformation of the A.G mispaired oligonucleotide d(CGCAAGCTGGCG). *Proc. Natl. Acad. Sci. USA* **1990**, 87, 6693-7.
15. Brown, T.; Hunter, W.N.; Kneale, G.; Kennard, O., Molecular structure of the G.A base pair in DNA and its implications for the mechanism of transversion mutations. *Proc. Natl. Acad. Sci. USA* **1986**, 83, 2402-6.
16. Greene, K. L.; Jones, R.L.; Li, Y.; Robinson, H.; Wang, A.H.; Zon, G.; Wilson, W.D., Solution structure of a GA mismatch DNA sequence, d(CCATGAATGG)2, determined by 2D NMR and structural refinement methods. *Biochemistry* **1994**, 33, 1053-62.
17. Pozhitkov, A.; Noble, P.A.; Domazet-Loso, T.; Nolte, A.W.; Sonnenberg, R.; Staehler, P.; Beier, M.; Tautz, D., Tests of rRNA hybridization to microarrays suggest that hybridization characteristics of oligonucleotide probes for species discrimination cannot be predicted. *Nucleic Acids Res.* **2006**, 34, e66.

18. Riccelli, P.V.; Merante, F.; Leung, K.T.; Bortolin, S.; Zastawny, R.L.; Janeczko, R.; Benight, A.S., Hybridization of single-stranded DNA targets to immobilized complementary DNA probes: comparison of hairpin versus linear capture probes. *Nucleic Acids Res.* **2001**, 4, 996-1004.
19. Meunier-Prest, R.; Raveau, S.; Finot, E.; Legay, G.; Cherkaoui-Malki, M.; Latruffe, N., Direct measurement of the melting temperature of supported DNA by electrochemical method. *Nucleic Acids Res.* **2003**, 31, e150.
20. Dunbar, SA., Applications of Luminex xMAP technology for rapid, high-throughput multiplexed nucleic acid detection. *Clin. Chim. Acta* **2006**, 363, 71-82.
21. Strachan, T.; Read, A.P., Human Molecular Genetics 2. *Bios* **1999**, John Wiley and Sons Inc., New York, (2ed).
22. Hutton, JR., Renaturation kinetics and thermal stability of DNA in aqueous solutions of formamide and urea. *Nucleic Acids Res.* **1977**, 10, 3537-55.
23. Naef, F.; Magnasco, M.O., Solving the riddle of the bright mismatches: labeling and effective binding in oligonucleotide arrays. *Phys. Rev. E. Stat. Nonlin. Soft Matter Phys.* **2003**, 68, 011906.

5. Enzyme accessibility on a suspension array

5.1 Introduction

There are numerous approaches to identifying single-nucleotide polymorphisms (SNPs) and diagnosing point mutations in DNA.[1] Despite the many methodologies available for DNA analysis, no technology for identifying SNPs and mutations has become a widely accepted standard, and the task of choosing a method is difficult.[2] Also, an estimated 20-30% of SNPs in large-scale SNP studies are rejected because of failure of the genotyping method.[3]

The available technologies today can be divided in two basic categories.[4] The first, based on direct hybridisation only, *e.g.*, in the format of microarrays, introduced in previous chapters, promises very high throughputs but with a limited robustness. The distinction between the sequence variants is based on small differences in thermal stability between mismatched and perfectly matched hybrids formed with the targeting probes. Microarrays are utilised mostly, at present, in gene expression analysis. [5] The other category concerns methodologies assisted by enzymatic activity in DNA analysis. These methodologies rely on the hybridisation of oligonucleotides at either the upstream region of the variation or hybridisation on the top of the sequence carrying the variation. Hybridisation at the upstream region requires a DNA polymerase for reading through the variable region and this is the basis of DNA sequencing techniques such as single base chain extension and pyrosequencing. Hybridisation on the variable region may require the use of DNA ligases or nucleases. Examples of these methodologies are padlock probes and digestion and/or ligation-mediated analysis.[6]

It is therefore of interest as well as to look at hybridisation on solid supports to look at the applicability of these methodologies for DNA analysis on beads and compare the performance of the enzymes in solution and on solid phase. The following sections introduce some of these enzyme-mediated processes of DNA analysis and particular

relevance will be given to solid phase single base chain extension and restriction endonuclease digestion as they are on the basis of the experimental work carried out for this chapter.

5.1.1 Solid phase single base chain extension

Solid phase single-nucleotide primer extension also called solid phase minisequencing, [1, 5, 7-11] or genetic bit analysis, [12, 13] was first introduced by Syvanen and co-workers.[14] It is one of the most accurate methods for detecting SNPs.[1, 5] Solid phase single base chain extension is more robust than methods based on direct hybridisation only, [5, 11] as the chemistry is based on the high accuracy of the nucleotide incorporation reaction catalysed by a DNA polymerase. The process is outlined in Figure 5.1.

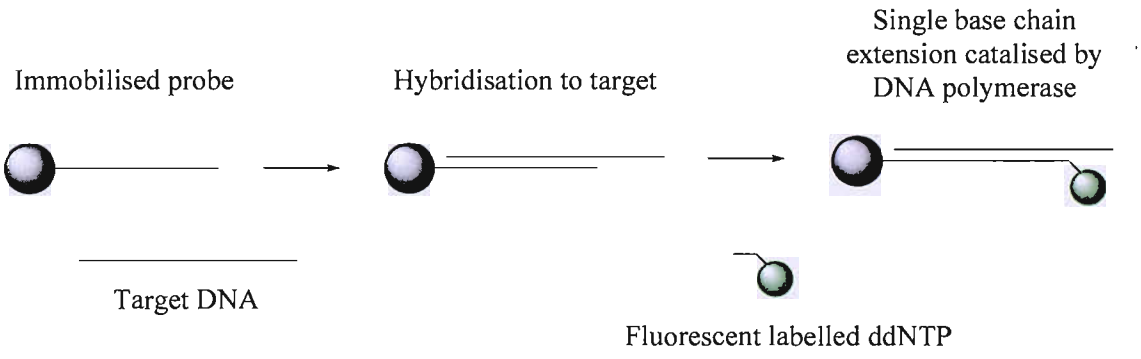


Figure 5.1 Detection of unlabelled DNA targets by single base chain extension.

In solid phase single base chain extension a DNA polymerase uses a tethered DNA probe as a primer in an extension reaction in which the substrate is a fluorescently labelled chain terminator ddNTP. [15, 16] The enzyme incorporates and extends only one base; that which is complementary to the next base in the target providing information on the sequence of the target.[11, 17] This method is particularly useful for identifying SNPs. By choosing a tethered primer adjacent to a polymorphic locus, extension with a fluorescent ddNTP, reveals the identity of the complementary base at the polymorphic

site.[11, 17-19] The platform can also be extended to multiplex SNPs analysis by interrogating targets with different loci variations in the presence of four differently labelled nucleotides. [4, 5, 11, 16, 18, 20]

5.1.1.1 DNA polymerases

DNA polymerases are responsible for DNA replication and for maintaining the integrity of genomes.[21] These enzymes were discovered by Kornberg and colleagues[22, 23] and constitute a critical component in DNA sequencing methods. DNA polymerases have the ability to elongate the DNA by adding either dNTPs or ddNTPs or even their fluorescently labelled analogues into a growing DNA chain (Figure 5.2). [21, 24-26]

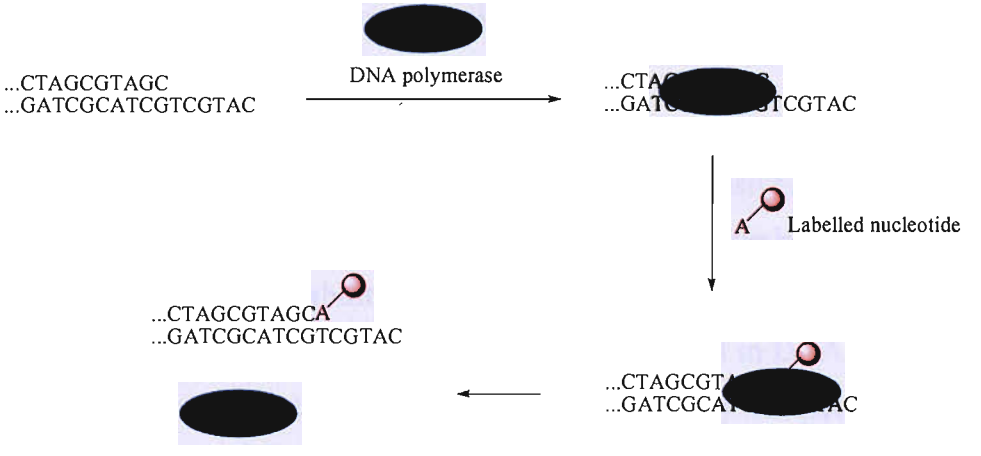


Figure 5.2 Incorporation of a labelled nucleotide by a DNA polymerase.

The ability to incorporate ddNTPs or dNTPs varies from enzyme to enzyme[15, 24-28] and is also influenced by the template.[29] This incorporation is remarkably affected if the nucleotide bears a fluorophore and most polymerases strongly discriminate against the incorporation of dye-labelled dNTPs and ddNTPs. [15, 21, 27, 28, 30] This has been attributed to the hydrophobic nature or bulkiness of the fluorescent dye groups and dye-enzyme interactions.[21, 30, 31] Moreover, the physico-chemical nature of the attached fluorophores

or other modifying groups and their linker moieties is also known to impart influence.[20, 21, 30, 31] So far a study addressing how efficiently or accurately DNA polymerases incorporate dye-labelled dNTPs and ddNTPs has not been fully addressed.[32]

DNA polymerase fidelity is the ability to discriminate between correct and incorrect nucleotide insertion[33] and is another factor that can affect the accuracy of SNPs genotyping by minisequencing.[32, 34] The precise structural features of the DNA substrate that are recognised and used by the polymerase to discriminate against the incorrect nucleotide insertion are complex. [23, 32, 35, 36]

The 3'-5' proofreading exonuclease activity is the ability of a polymerase to remove a misincorporated nucleotide in the newly synthesised DNA. The lack of 3'-5' proofreading means the enzyme does not have that ability and might generate errors in sequencing. On the other hand, a strong 3'-5' exonuclease activity might also hinder the ability of the polymerase to incorporate nucleotide analogues as the proofreading ability results in removal of the added labelled nucleotide. [21, 37]

Taq DNA polymerase is widely known and has been extensively used in PCRs to amplify DNA because it is fast and has a high turnover number,[25, 38] it is highly processive,[38] exhibits thermostability,[25, 26, 38] and is inexpensive. Due to these properties, *Taq* has also been one of the most predominant enzymes used in DNA sequencing and thus has been extensively studied with regard to its ability to incorporate dNTPs, ddNTPs and dye-labelled ddNTPs.[28] *Taq* has no detectable 3'-5' proofreading exonuclease.[25, 26, 28, 39, 40] However, according to Eckert and Kunkel, the fidelity of *Taq* can be greatly improved by optimising the concentrations of $MgCl_2$ and dNTPs in extension reactions.[40] *Taq* is also known for strongly discriminating against the incorporation of ddNTPs, [25, 26] specially if these bear a fluorescein label.[15]

The discrimination against the incorporation of ddNTPs might be overcome with the use of Thermo Sequenase™ DNA polymerase. Thermo Sequenase™ is a mutant derived from *Taq*, also lacking 3'-5' proofreading exonuclease,[29] and favours the incorporation of ddNTPs over dNTPs. [1, 24, 26, 32] According to Tabor and co-workers,[26]

Thermo Sequenase™ incorporates fluorescein-ddNTPs approximately three orders of magnitude more efficiently than *Taq*.

If 3'-5' proofreading exonuclease is required, *Pfu* and *Vent* DNA polymerases are amongst the most commonly used enzymes, however these show difficulties in incorporating modified ddNTPs, particularly those labelled with fluorophores.[27, 28]
The ability of a DNA polymerase to discriminate between dNTP and ddNTP can be altered dramatically by substitution of a single amino acid in a conserved region of the polymerase.[24-29] This is highly promising and useful as efficient ddNTP usage provides several advantages including more uniform signal intensities, reduced costs, and lower background associated with using lower dye-terminator concentrations in fluorescent DNA sequencing.[32]

5.1.2 Pyrosequencing

Pyrosequencing[3, 6, 41] is a method of DNA sequencing developed by Ronaghi and co-workers[42] in the late 1990s. It is based on a chemiluminescent enzymatic reaction, which is triggered when a molecular recognition event occurs. Essentially, the method allows sequencing of a single strand of DNA by synthesising the complementary strand along it. Each time a nucleotide is incorporated into the growing chain a cascade of enzymatic reactions is triggered which results in a light signal.[6, 41] Figure 5.3 outlines the pyrosequencing cycle.

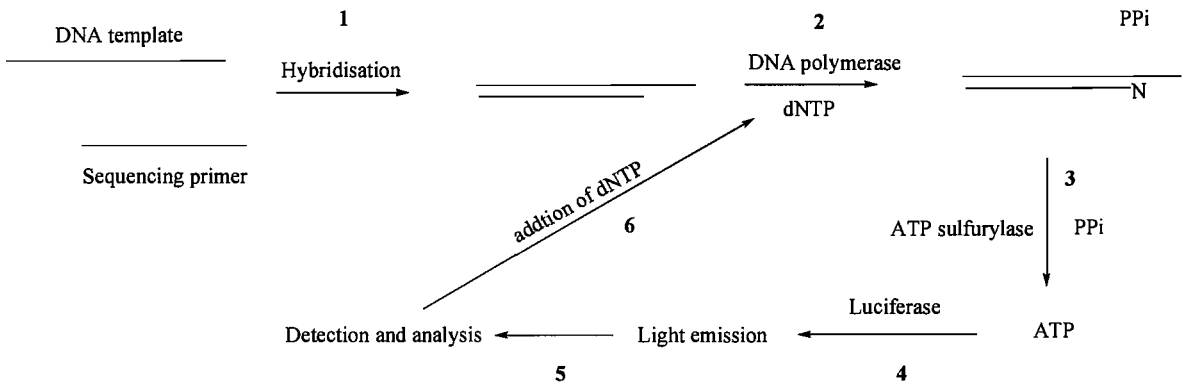


Figure 5.3 Outline of the pyrosequencing methodology.

A sequencing primer is hybridised to a single stranded, PCR amplified, DNA template (**1**), and incubated with the enzymes, DNA polymerase, ATP sulfurylase, luciferase and apyrase, and the substrates, adenosine 5' phosphosulfate and luciferin. The first of four deoxynucleotide triphosphates is added to the reaction. DNA polymerase catalyses the incorporation of the dNTP into the DNA strand, if it is complementary to the base in the DNA template. Each incorporation event is accompanied by release of pyrophosphate (PPi) in a quantity equimolar to the amount of incorporated nucleotide (**2**). ATP sulfurylase quantitatively converts PPi to ATP in the presence of adenosine 5′ phosphosulfate (**3**). This ATP drives the luciferase-mediated conversion of luciferin to oxyluciferin that generates visible light in amounts that are proportional to the amount of ATP (**4**). A pyrogram is obtained in which the light produced is proportional to the number of nucleotides incorporated (**5**). Apyrase, a nucleotide degrading enzyme, continuously degrades unincorporated dNTPs and excess ATP. When degradation is complete, another dNTP is added and the pyrosequencing cycle starts again (**6**) providing information about the DNA sequence.

Pyrosequencing provides the same sequencing accuracy as conventional sequencing methods being highly versatile for SNP genotyping and SNP discovery.[41] It also possesses the advantages of dispensing with the need for labelled nucleotides, labelled primers and electrophoresis and a large number of samples can be analysed in a short period of time.

It is also a technique amenable to be transferred to solid phase. Ronaghi *et al* [6, 41] report on the solid phase preparation of the DNA templates by immobilising biotinylated PCR products onto streptavidin coated beads. Solid phase strategies allow for cheaper and easier template preparation which is more adaptable for automation. A solid phase approach can also be multiplexed which enables multiple parallel analysis of one or several target DNA sequences thereby reducing the cost and time consumption in the preparation of samples and genotyping. [2, 3, 6, 41, 43]

5.1.3 Padlock probes

Padlock probes[44, 45] are linear oligonucleotide molecules designed to have end sequences that are complementary to a sequence to be detected in a target DNA molecule. These linear oligonucleotides are typically between 90 and 100 bases long,[46, 47] and they are designed to hybridise end to end on the target sequence. The variant base in the target sequence is positioned at the 3' end of the probe. If there is a perfect match between the probe and the target sequence, a DNA ligase can join the ends, turning the probe into a circle, while mismatched 3' ends will inhibit ligation. In this manner single-nucleotide substitutions can be discriminated.[44-46] Figure 5.4 outlines the processes.

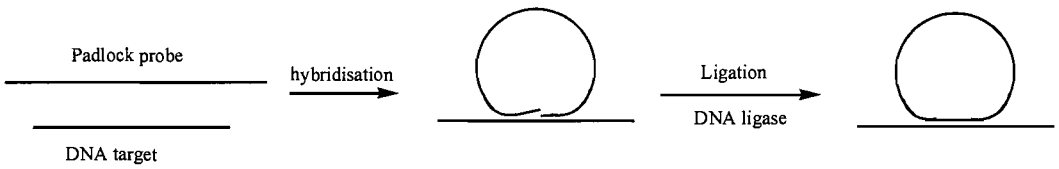


Figure 5.4 Outline of a padlock probe hybridisation and ligation processes.

A circular molecule is formed that can be very specifically detected using any DNA amplification method but, being a circular molecule, it is particularly well-suited to be amplified by rolling circle amplification (Figure 5.5).[44, 48, 49]

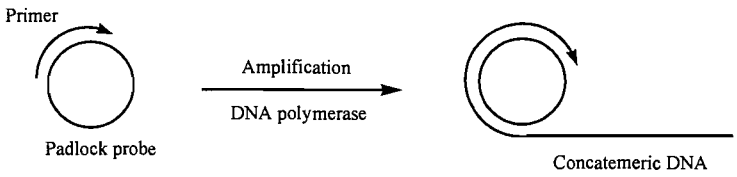


Figure 5.5 The principle of linear rolling circle amplification.

The padlock strategy also offers the advantage of multiplexing. A very large number of probes can be analysed simultaneously in a sample since any cross reactivity between

probes will generate linear molecules that will be removed by exonuclease treatment.[47] The result from a padlock assay can be analysed in a simple plate reader or in several other formats including microarrays.[47, 50]

5.1.4 Restriction endonuclease and ligation mediated strategies

Restriction endonucleases are useful tools to generate DNA fragments with different lengths and are in the forefront of the enzymes used in genetic engineering. [51, 52] A restriction endonuclease functions by "scanning" the length of a DNA molecule. Once it encounters its particular specific recognition sequence, it will bond to the DNA molecule and will make one cut in each of the two sugar-phosphate backbones of the double helix. The positions of these two cuts, both in relation to each other, and to the recognition sequence itself, are determined by the identity of the restriction endonuclease used to cleave the molecule in the first place.

Different endonucleases yield different sets of cuts, but one endonuclease will always cut a particular base sequence the same way, no matter what DNA molecule it is acting on.[53] Figure 5.6 outlines the endonuclease process for a palindromic cleavage generating overhang products.

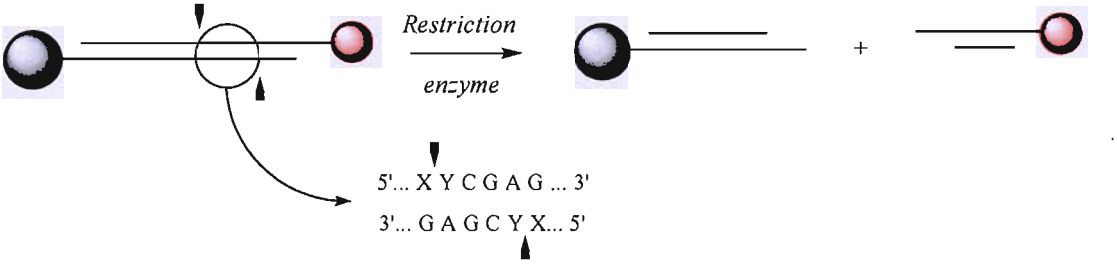


Figure 5.6 Restriction recognition site and endonuclease cleavage generating overhang products.

Not all restriction endonucleases produce overhang[51, 54, 55] products like outlined in Figure 5.6. Many endonucleases cleave the DNA backbones in positions directly opposite each other producing blunt ends.[52]

Overhangs can also be further used in subsequent ligation reactions to complementary overhang sequences from different DNA sources, bacterial for example, originating recombinant DNA molecules.[53-55] Brenner and co-workers[55-57] report on a sequencing strategy employing successive cycles of ligation and restriction digestion on immobilised DNA duplexes. Figure 5.7 outlines the cyclic strategy. Adaptors encoded with a specific sequence are ligated to the ends of the DNA molecules attached to the beads. A fluorescently labelled decoder is then hybridised to the adaptor ends thereby providing information on the sequence of the last four bases. The encoded adapter is then removed by digestion with an endonuclease exposing the next four nucleotides as a single stranded overhang and ready to undergo a subsequent interrogation cycle.

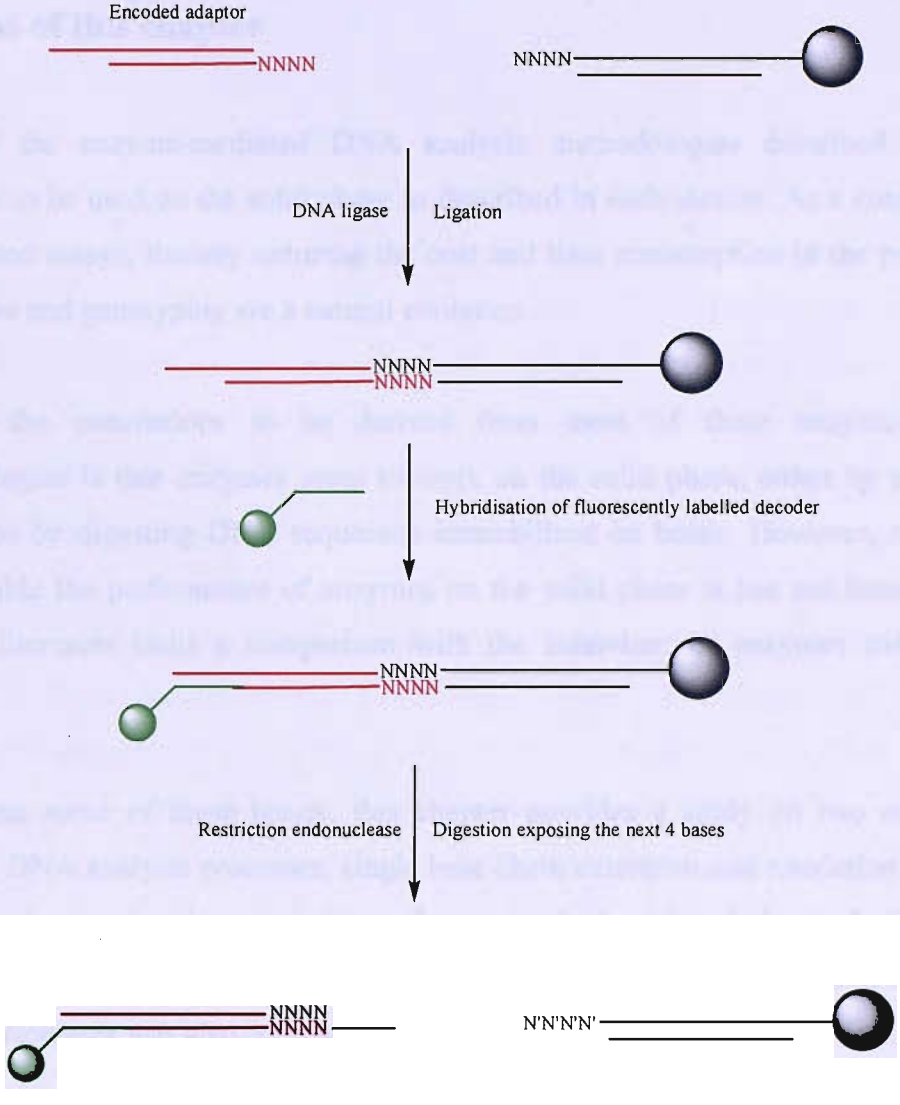


Figure 5.7 Sequencing strategy composed of successive cycles of ligation and digestion on DNA duplexes immobilised on beads as described by Brenner and co-workers.[55]

Restriction overhang products can also be used in subsequent chain extension reactions by DNA polymerases.[51] Restriction enzymes can also be used to find specific SNPs.[51, 58-62] The principle behind this approach relies in an appropriate restriction endonuclease whose recognition sequence has been altered or introduced by the SNP in order to recognise it.[61]

5.2 Aims of this chapter

Most of the enzyme-mediated DNA analysis methodologies described here are amenable to be used on the solid phase as described in each section. As a consequence, multiplexed assays, thereby reducing the cost and time consumption in the preparation of samples and genotyping are a natural evolution.

One of the conclusions to be derived from most of these enzyme-mediated methodologies is that enzymes seem to work on the solid phase, either by extending, ligating or by digesting DNA sequences immobilised on beads. However, a study of how reliable the performance of enzymes on the solid phase is has not been reported and the literature lacks a comparison with the behaviour of enzymes exhibited in solution.

To address some of these issues, this chapter provides a study on two enzymatic-mediated DNA analysis processes, single base chain extension and restriction digestion activity and compares the enzymatic performance obtained in solution and solid phase. The incorporation and misincorporation of fluorescently labelled dNTP and ddNTP by DNA polymerases was also compared in solution and solid phase.

5.3 Results and discussion

The detection of unlabelled target DNA by single base chain extension and restriction endonuclease digestion experiments were already introduced in Chapters 1 and 2. The lack of success attained in the preliminary base extension experiments carried out in Chapter 2, was first attributed to the extensive non-specific binding of the labelled nucleotide to the beads and second to the ineffective EDC mediated chemistry of probe attachment which rendered the DNA construct upon hybridisation, less available for enzyme accessibility. [63-65] The lack of success of endonuclease digestion was also attributed to the poor method of probe attachment.

However, the introduction of the avidin-biotin system of probe attachment,[7, 66-68] described in Chapters 3 and 4, led to a high improvement in hybridisation efficiency and hybridisation discrimination. On the basis of that success, single base chain extension experiments were resumed and the incorporation of fluorescent ddNTPs and dNTPs were attempted. Restriction endonuclease digestions were also continued. The set of experiments described in this chapter aimed at a comparison between yield, rate and specificity of enzyme activity in solution and on solid phase.

5.3.1 Single base chain extension

5.3.1.1 Fluorescent nucleotides used for single base extension reactions

Fluorescein-ddATP (Figure 5.8) and Cy5-dUTP (Figure 5.9) were chosen to compare and assess the extent of incorporation of a ddNTP versus a dNTP by a DNA polymerase onto a DNA template in solution and on solid phase.

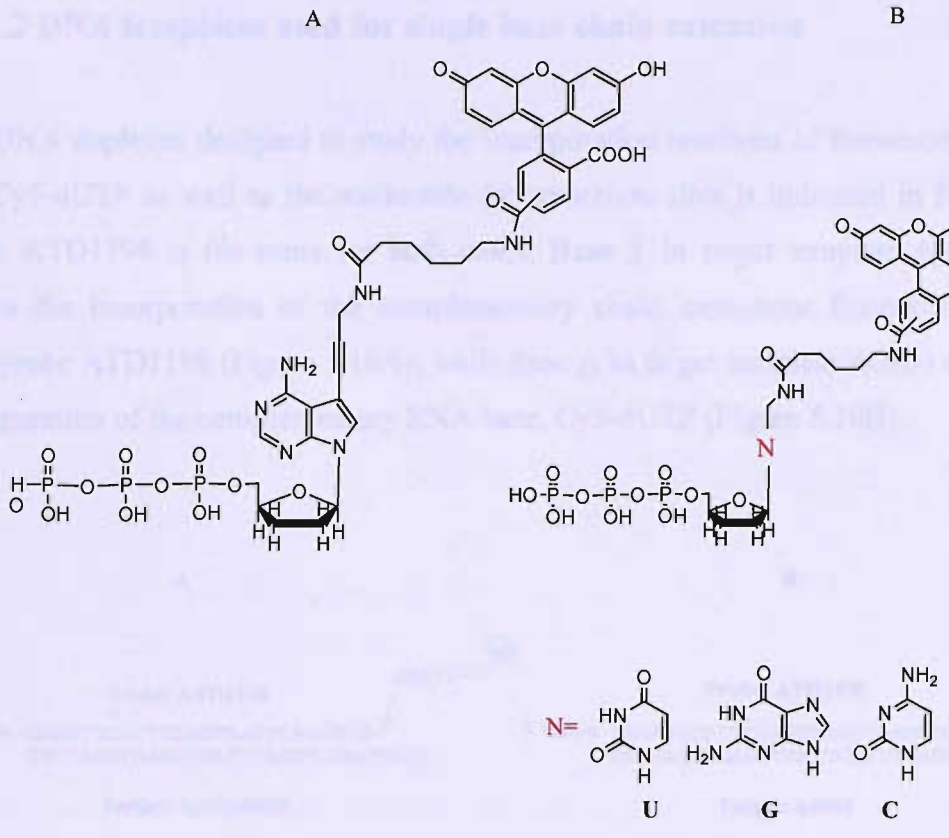


Figure 5.8 A – Chemical structure of fluorescein-ddATP. B – Chemical structure of fluorescein-ddNTPs, N = U, G and C used in misincorporation discrimination experiments.

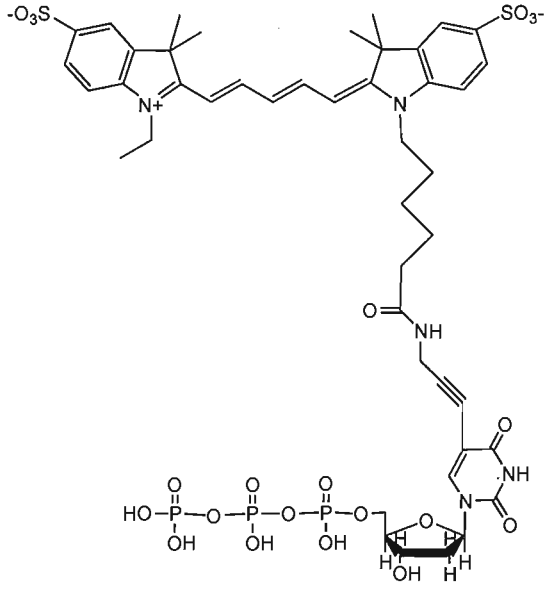


Figure 5.9 Chemical structure of Cy5-dUTP.

5.3.1.2 DNA templates used for single base chain extension

The DNA duplexes designed to study the incorporation reactions of fluorescein-ddATP and Cy5-dUTP as well as the nucleotide incorporation sites is indicated in Figure 5.5. Probe ATD1198 is the same for both cases. Base T in target template 4goligo0015 directs the incorporation of the complementary chain terminator fluorescein-ddATP onto probe ATD1198 (Figure 5.10A), while base A in target template A0960 directs the incorporation of the complementary RNA base, Cy5-dUTP (Figure 5.10B).

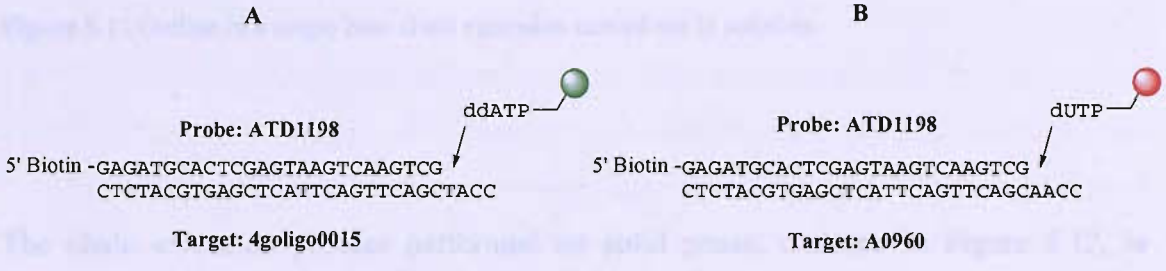


Figure 5.10 Oligonucleotide probe and targets sequences required for the base extensions and the indicated nucleotide incorporation site. A – template 4goligo0015 directs the incorporation of chain terminator ddATP onto probe ATD1198 and conversely, B – template A0960 directs the incorporation of base dUTP.

5.3.1.3 Performing base extensions in solution and on solid phase

Figure 5.11 outlines the single base chain extension experimental procedure carried out in solution using the DNA construct designed to accommodate the incorporation of fluorescein-ddATP as an example. The base extension in solution involved the incubation of both DNA probe and target template with enzyme and fluorescein-ddATP. The extended labelled duplexes were subsequently immobilised onto avidin coated GMA beads, through the 5' biotin group located on the probe. The beads were thoroughly washed to remove any non-specific bound labelled nucleotide and analysed by FACS.

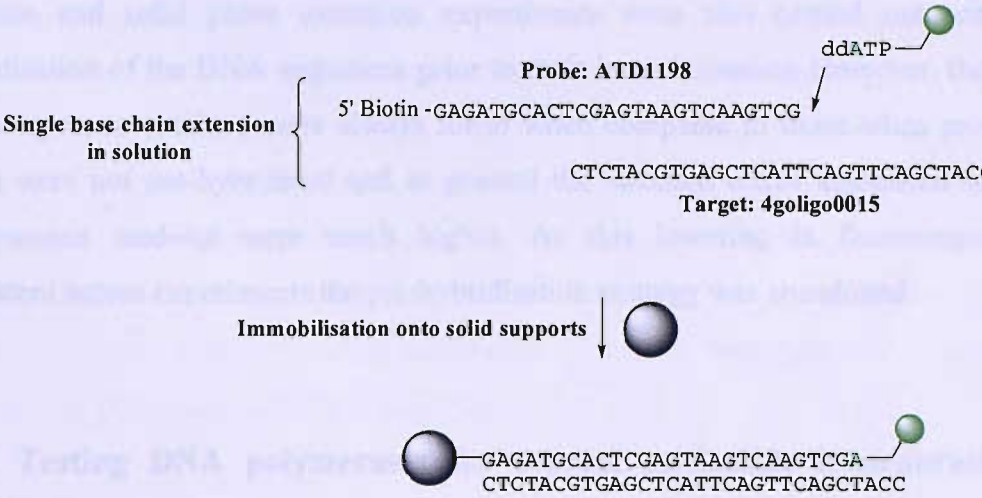


Figure 5.11 Outline of a single base chain extension carried out in solution.

The chain extension process performed on solid phase, outlined in Figure 5.12, is similar to that described for solution, only the DNA probe was immobilised onto the solid supports beforehand.[1] The incubation reaction proceeded with agitation[69, 70] to ensure the beads would not settle out of solution. After chain extension the beads were washed and analysed by FACS.

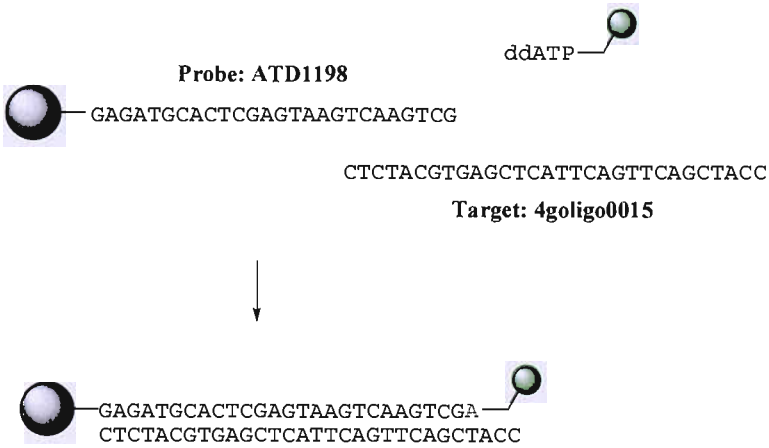


Figure 5.12 Outline of a single base chain extension carried out on solid phase.

Solution and solid phase extension experiments were also carried out with pre-hybridisation of the DNA sequences prior to their immobilisation. However, the levels of fluorescence obtained were always lower when compared to those when probe and target were not pre-hybridised and in general the standard errors associated with the fluorescence read-out were much higher. As this lowering in fluorescence was consistent across experiments the pre-hybridisation strategy was abandoned.

5.3.2 Testing DNA polymerases for correct nucleotide incorporation in solution

For the preliminary base extension experiments, three different DNA polymerases; *Taq*, *Vent* and *Pfu Turbo*, were tested to evaluate their ability to incorporate a nucleotide onto a DNA construct. Moreover, the experiment aimed to assess which enzyme possessed the ability to discriminate between correct versus incorrect fluorescein-ddNTP incorporation. Each DNA polymerase was incubated with the DNA construct displayed in Figure 5.10A in the presence of fluorescein-ddATP. As a control for monitoring nucleotide misincorporation, base extension experiments in the presence of fluorescein-ddUTP (Figure 5.8B) were also carried out. Figure 5.13 shows the fluorescence results exhibited by the beads obtained for the different DNA polymerases used.

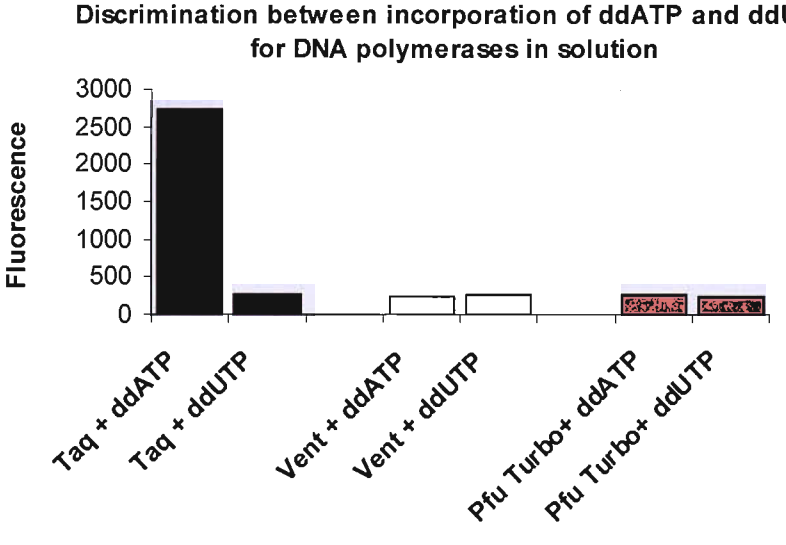


Figure 5.13 Fluorescence (± SE) exhibited by the incubation of the DNA duplexes with the respective enzymes and fluorescein-ddNTPs.

Results show that *Taq* was the only enzyme that successfully incorporated fluorescein-ddATP (2,733±9.3). The fluorescence exhibited by the incubation of *Taq* with fluorescein-ddUTP (255±1.3) is very close to that of a control consisting of avidin coated GMA beads incubated with fluorescein-ddATP and *Taq* (232±1.3) indicating undetectable levels of misincorporation. The fact that *Taq* is known for discriminating strongly against the incorporation of ddNTPs[24-28] and that it only poorly incorporates those fluorescein labelled[15] might explain the not very high level of fluorescence observed for fluorescein-ddATP incorporation.

Vent and *Pfu Turbo* failed to incorporate fluorescein-ddATP, as the exhibited low levels of fluorescence comparable to that of the control indicated. This is not surprising as these two enzymes have been reported as having strong difficulties incorporating modified nucleotides, such as ddNTPs, especially those labelled with fluorophores.[27, 28]

In order to determine the incorporation/misincorporation response of *Taq* to the full set of labelled dideoxynucleotides, the extension experiments were extended to fluorescein-ddCTP and fluorescein-ddGTP (Figure 5.14).

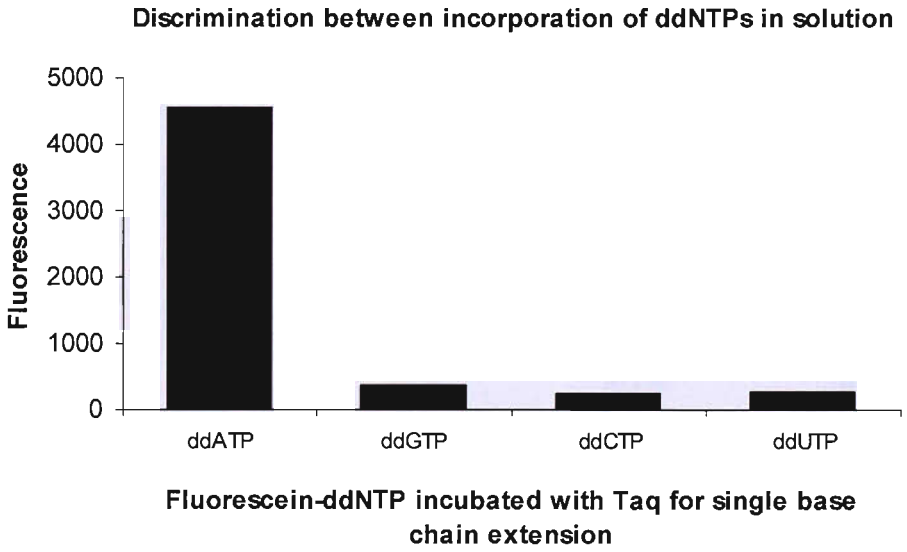


Figure 5.14 Fluorescence (± SE) exhibited by the incubation of the DNA duplexes with *Taq* and the fluorescein-ddNTPs.

Fluorescence results are consistent in that *Taq* is capable of discriminating against fluorescein-ddNTP misincorporation. The fluorescence values exhibited by the incubation with ddCTP (239±1.5), ddUTP (286±1.6) and ddGTP (355±1.6) are very close to that of a control consisting of avidin coated GMA beads incubated with fluorescein-ddATP and *Taq* (225±1.3) indicating undetectable levels of misincorporation for these nucleotides.

5.3.3 Single base chain extension on solid phase

Previous experiments showed that *Taq* successfully incorporated fluorescein-ddATP in solution. To determine the ability of *Taq* to reproduce the incorporation process on solid phase, GMA beads with immobilised DNA probe were incubated with target template, enzyme and fluorescein-ddATP, according to Figure 5.12. The experiment was also carried out in the presence of fluorescein-ddUTP to assess the level of misincorporation on the solid phase.

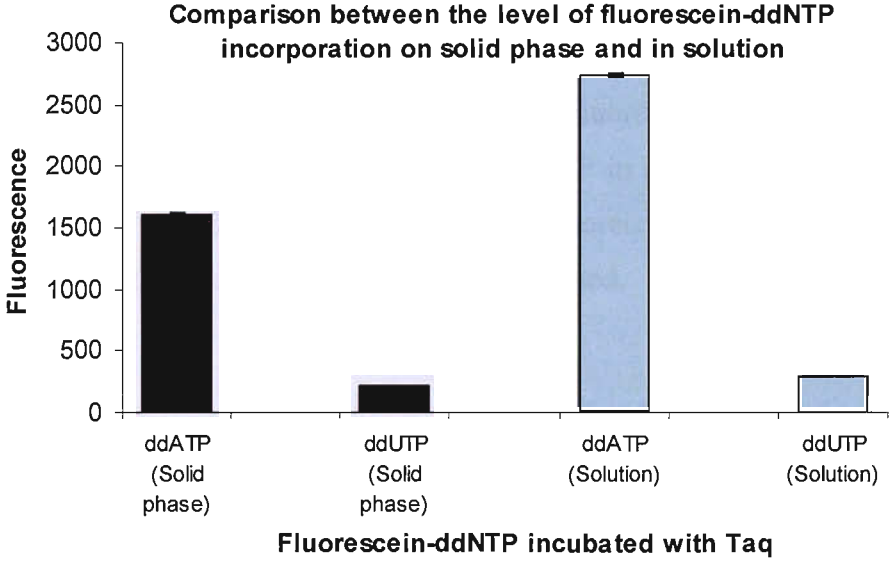


Figure 5.15 Level of fluorescein-ddNTP incorporation by the chain extension experiments on solid phase and in solution.

Figure 5.15 shows the fluorescence values obtained for the chain extension experiments on solid phase and in solution. Results demonstrate that the level of fluorescein-ddATP incorporation is much lower on the solid phase. This is not unsurprising, as the presence of the solid supports to which the DNA probes are attached constraint the way in which the enzyme interacts with the substrates. DNA is not so free to diffuse as it would in solution, and reduces the reaction rates and leads to less effective ddNTP incorporation. This fact combined with the strong discrimination *Taq* displays against fluorescein-ddNTPs incorporation[15, 24-28] explains the lower fluorescence observed for the solid phase assays.

Results also show that discrimination against misincorporation is achieved in both solution (294±1.2) and solid phase (245±0.9) as demonstrated by the low fluorescence of the samples incubated with fluorescein-ddUTP. A control consisting of immobilised DNA probe incubated with fluorescein-ddATP and *Taq* in the absence of target template yielded a fluorescence of 248±1.1, indicating the absence of single strand extension.

As already mention in section 5.3.1.3, solution and solid phase extension experiments were also carried out with pre-hybridisation of the DNA sequences, but the levels of fluorescence obtained were always lower when compared to those when probe and target were not pre-hybridised. As an example, the fluorescence obtained in a given experiment for the incorporation of fluorescein-ddATP in solution was 1,321±5.5 and on solid phase was 430±6.8. As this lowering in fluorescence was consistent across experiments the pre-hybridisation strategy was abandoned.

In order to improve the incorporation of ddNTPs on solid phase the chain extension experiment was carried out with Thermo Sequenase™ DNA polymerase.[29] This polymerase is known for its excellent sequencing ability[24] and increased capacity for ddNTPs incorporation when compared to *Taq*. [1, 24, 26, 28] As expected, Figure 5.16 shows the improved capacity of Thermo Sequenase™ for fluorescein-ddATP incorporation (8,253±31.1) when compared to that of *Taq* (1,587±3.7).

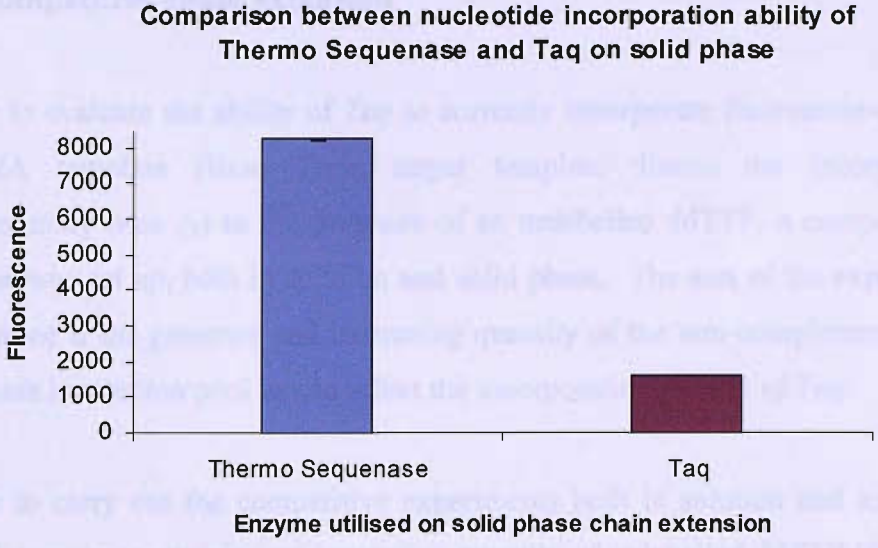


Figure 5.16 Comparison between the level of fluorescein-ddATP incorporation on solid phase by Thermo Sequenase™ and *Taq*.

5.3.4 Competitive chain extension

In order to evaluate the ability of *Taq* to correctly incorporate fluorescein-ddATP onto the DNA template (Base T in target template directs the incorporation of complementary base A) in the presence of an unlabelled ddTTP, a competitive chain extension was set up, both in solution and solid phase. The aim of the experiment was to determine if the presence and increasing quantity of the non-complementary ddTTP in the same incubation pool would affect the incorporation fidelity of *Taq*.

In order to carry out the competitive experiments both in solution and solid phase, a series of samples containing an increasing quantity of unlabelled ddTTP (0, 10, 20, 40, 100 and 200 µM) and a lower and constant quantity of fluorescein-ddATP (2 µM) were incubated with the DNA and *Taq*. Figure 5.17 illustrates the process by outlining the solution experiments.

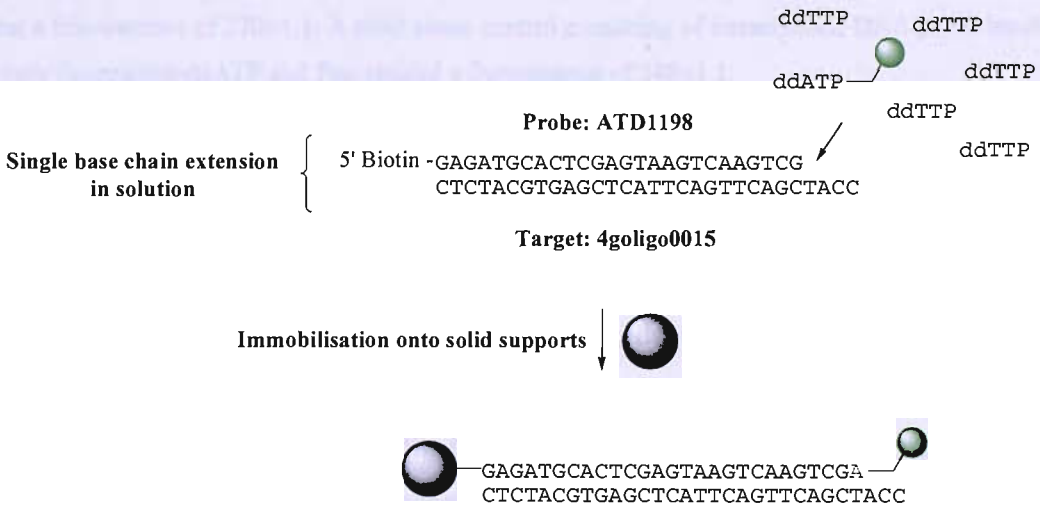


Figure 5.17 Outline of the competitive chain extension experiments in solution to determine if *Taq* discriminated correctly against the incorporation of unlabelled ddNTP in the same incubation pool.

Figure 5.18 shows the fluorescence results obtained for solution and solid phase extensions with increasing concentration of ddTTP.

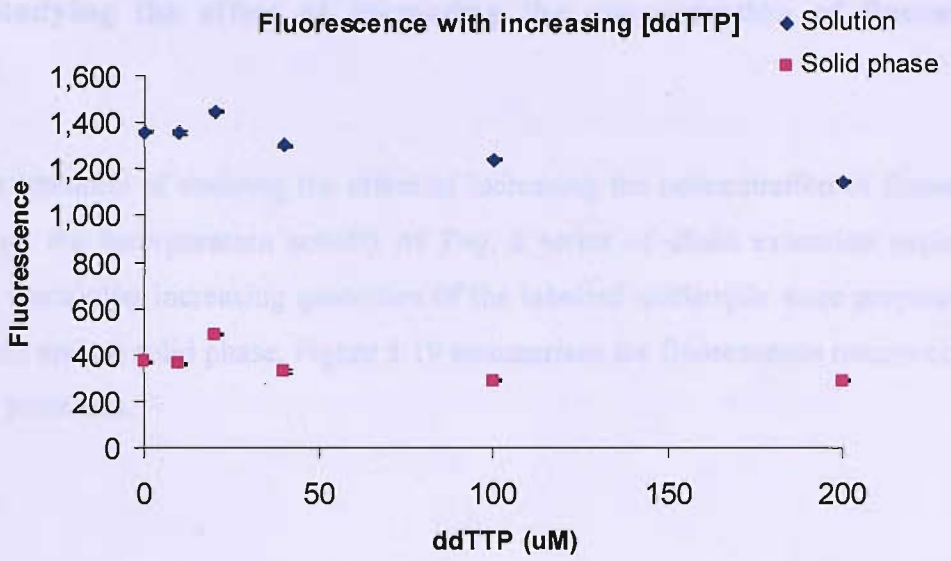


Figure 5.18 Fluorescence (± SE) of the beads for the competitive chain extension carried out in solution and solid phase. The concentration of fluorescein-ddATP was kept constant in all samples at 2 μM. A solution control containing 200 μM ddTTP and no fluorescein-ddATP yielded a fluorescence of 254±0.8 and a control consisting of avidin coated GMA beads incubated with only fluorescein-ddATP and *Taq* to assess the level of non-specific binding of the fluorescent nucleotide at the specified concentration yielded a fluorescence of 270±1.1. A solid phase control consisting of immobilised DNA probe incubated with only fluorescein-ddATP and *Taq* yielded a fluorescence of 248±1.1.

The fluorescence obtained for the competitive chain extension carried out in solution showed higher fluorescence as expected. Both in solution and on solid phase, the fluorescence of the beads followed the same trend, decreasing slightly towards higher concentrations of unlabelled ddTTP (from 40 to 200 μM). A fact not understood was that the fluorescence obtained in both processes, for the samples incubated with 20 μM ddTTP exhibited the highest fluorescence. The increasing concentration of ddTTP, at the working concentration of fluorescein-ddATP (2 μM) did not seem to have a significant impact on the incorporation of the labelled nucleotide, and overall *Taq* was considered to be selective in discriminating against ddTTP misincorporation.

5.3.5 Studying the effect of increasing the concentration of fluorescein-ddATP

With the intention of studying the effect of increasing the concentration of fluorescein-ddATP on the incorporation activity of *Taq*, a series of chain extension experiment samples containing increasing quantities of the labelled nucleotide were prepared both in solution and on solid phase. Figure 5.19 summarises the fluorescence results obtained for both processes.

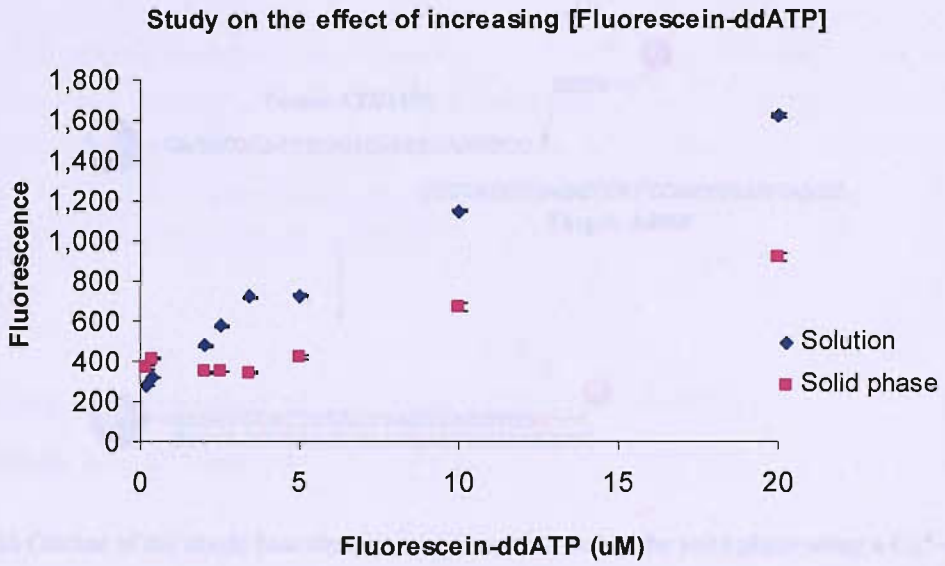


Figure 5.19 Fluorescence (± SE) of the beads on the study on the effect of increasing the concentration of fluorescein-ddATP in solution and on solid phase.

As expected, fluorescence results show that increasing the concentration of fluorescein-ddATP results in an increase on the incorporation of the nucleotide, both in solution and on solid phase. As observed in previous experiments, the level of nucleotide incorporation is higher in solution than on solid phase.

5.3.6 Incorporation of Cy5-dUTP in solution and on solid phase

Taq incorporates dNTPs at a rate that is several hundred to several thousands times that of ddNTPS.[26, 29, 38] Amongst the labelled nucleotides, Cy5-dUTP is one of the most commonly used as it possesses high incorporation efficiency, a good photostability and yield.[20, 71-73] Solution and solid phase chain extensions were set up, to evaluate and compare the level of Cy5-dUTP incorporation of *Taq*. Figure 5.20 outlines the solid phase procedure as an example, indicating the DNA sequences employed and the target templated addition of the labelled nucleotide.

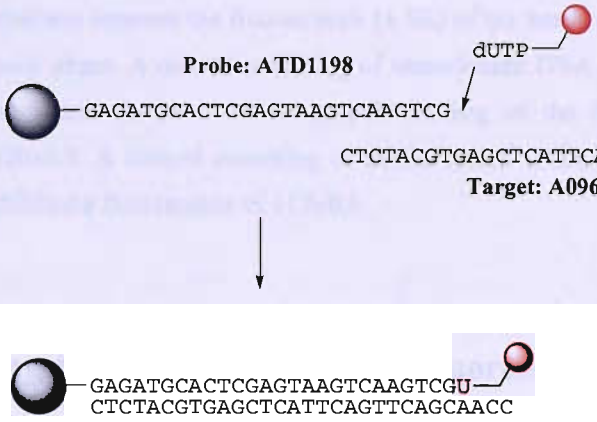


Figure 5.20 Outline of the single base chain extension carried out on the solid phase using a Cy5-dUTP.

As expected, the fluorescence values summarised in Figure 5.21 show that the incorporation is more effective in solution (5,777±28.6) than it is on solid phase (2,346±16.7). The control containing avidin coated GMA beads incubated with only Cy5-dUTP and *Taq* and the control consisting of immobilised DNA probe incubated with only Cy5-dUTP and *Taq* yielded fluorescences of 117±0.9 and 120±0.9 respectively. These low levels of fluorescence exhibited by the controls indicate undetectable levels of non-specific binding of the labelled nucleotide and also the absence of single strand extension.

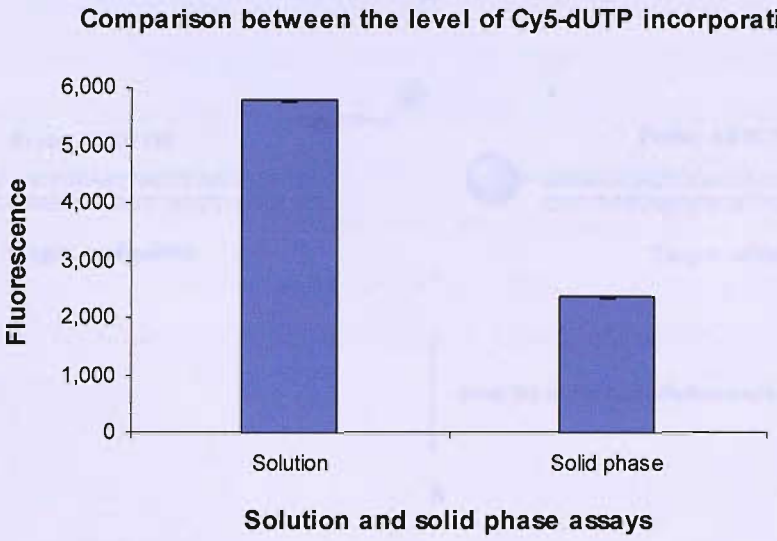


Figure 5.21 Comparison between the fluorescence (± SE) of the beads after Cy5-dUTP incorporation in solution and on solid phase. A control consisting of immobilised DNA probe incubated with only Cy5-dUTP and *Taq* to assess the level of non-specific binding of the fluorescent nucleotide yielded a fluorescence of 120±0.9. A control consisting of avidin coated GMA beads incubated with only Cy5-dUTP and *Taq* exhibited a fluorescence of 117±0.9.

5.3.7 Analysing the misincorporation of fluorescein-ddATP and Cy5-dUTP by changing target template

According to Figure 5.22A and as already explained in previous sections, base T in target 4goligo0015 acts as the template for the incorporation of complementary base A (fluorescein-ddATP) and conversely, base A in target A0960 acts as the template for the incorporation of complementary RNA base U.

Previous experiments showed that the level of selectivity of incorporation of fluorescein-ddATP was high. In order to assess the selectivity of incorporation of Cy5-dUTP, the formation of the non-complementary base pair T:Cy5-dUTP was investigated. In other words, the addition of the labelled nucleotides to the chain extension reactions was swapped (Figure 5.22B). Fluorescein-ddATP was incubated with target A0960 (A:fluorescein-ddATP) and Cy5-dUTP was incubated with target 4goligo0015 (T:Cy5-dUTP). Figure 5.23 summarises the fluorescence obtained for the solid phase chain extension experiments.

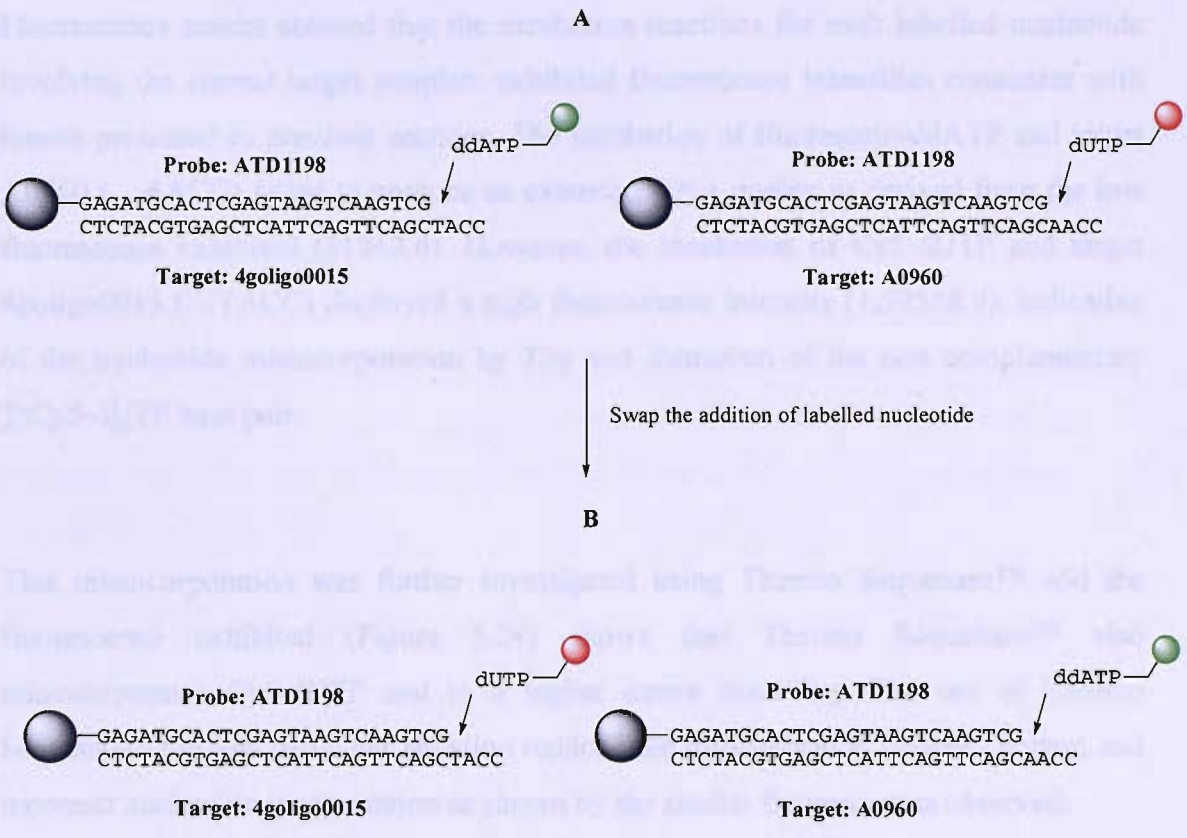


Figure 5.22 A- Oligonucleotide probe and targets sequences required for the respective base extensions and the indicated nucleotide incorporation site. B- Swapping the labelled nucleotides in the incubation samples.

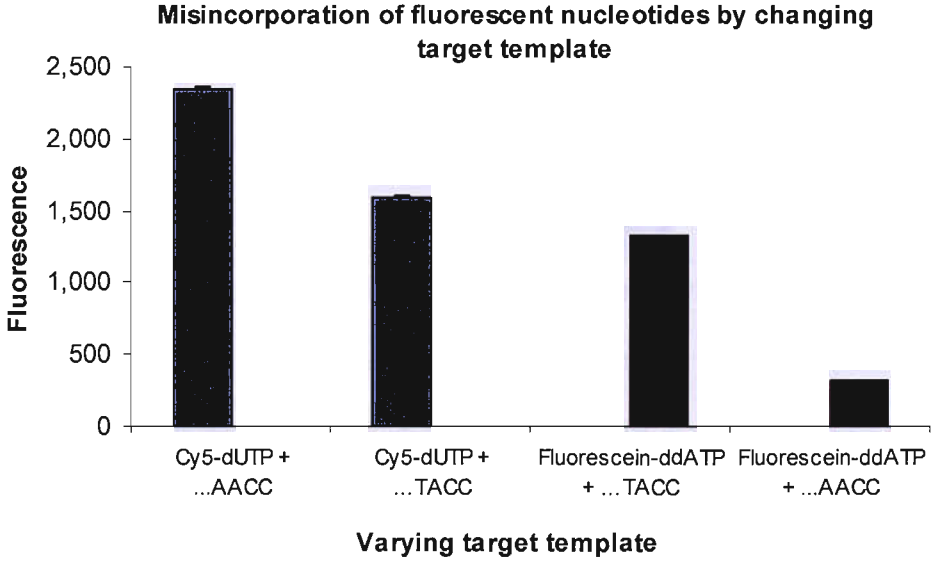


Figure 5.23 Comparison of the fluorescence (± SE) obtained for the incubation of Cy5-dUTP and fluorescein-ddATP with different target templates.

Fluorescence results showed that the incubation reactions for each labelled nucleotide involving the correct target template exhibited fluorescence intensities consistent with results presented in previous sections. The incubation of fluorescein-ddATP and target A0960 (...AACC) failed to produce an extended DNA duplex as derived from the low fluorescence exhibited (317±3.6). However, the incubation of Cy5-dUTP and target 4goligo0015 (...TACC) displayed a high fluorescence intensity (1,595±8.9), indicative of the nucleotide misincorporation by *Taq* and formation of the non complementary T:Cy5-dUTP base pair.

This misincorporation was further investigated using Thermo Sequenase™ and the fluorescence exhibited (Figure 5.24) shows that Thermo Sequenase™ also misincorporates Cy5-dUTP and to a higher extent than *Taq*. The use of Thermo Sequenase™ in this particular situation reduced the discrimination between correct and incorrect nucleotide incorporation as shown by the similar fluorescences observed.

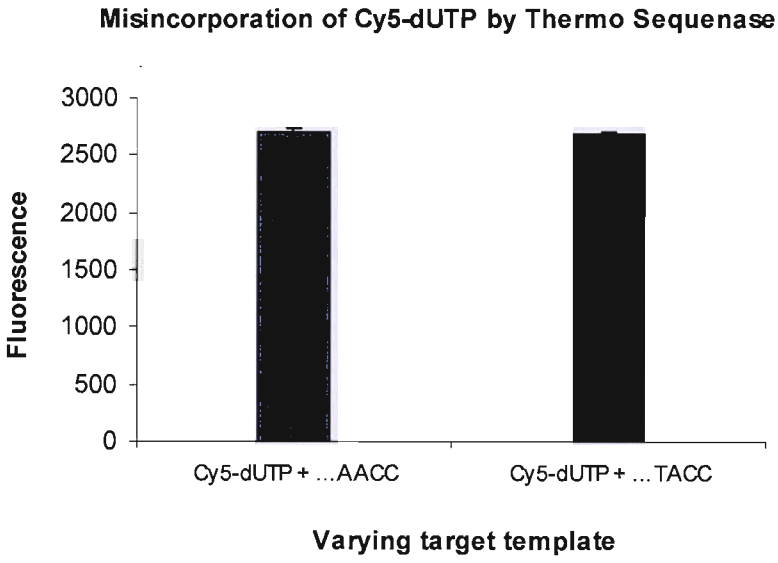


Figure 5.24 Comparison of the fluorescence (± SE) obtained for the incorporation of Cy5-dUTP with targets A0960 and 4goligo0015 using Thermo Sequenase™.

5.3.8 Solution and solid phase non-templated addition by *Taq*

Previous experiments showed low levels of selectivity of incorporation of Cy5-dUTP. To understand the nature of this non-specific labelling of Cy5-dUTP a chain extension reaction involving a blunt DNA construct was set up (Figure 5.25). Probe ATD1198 and target 4goligo26 form a blunt end duplex upon hybridisation, and the lack of a template overhang cannot provide coding information for the incorporation of either fluorescein-ddATP or Cy5-dUTP. Figure 5.26 summarises the fluorescence exhibited by the solution and solid phase blunt-end addition reactions involving fluorescein-ddATP and Cy5-dUTP.

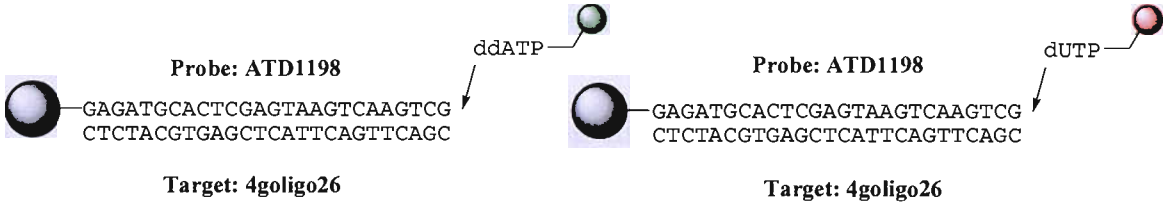


Figure 5.25 Substrate used to monitor the blunt end addition reaction of the labelled nucleotides on the solid phase.

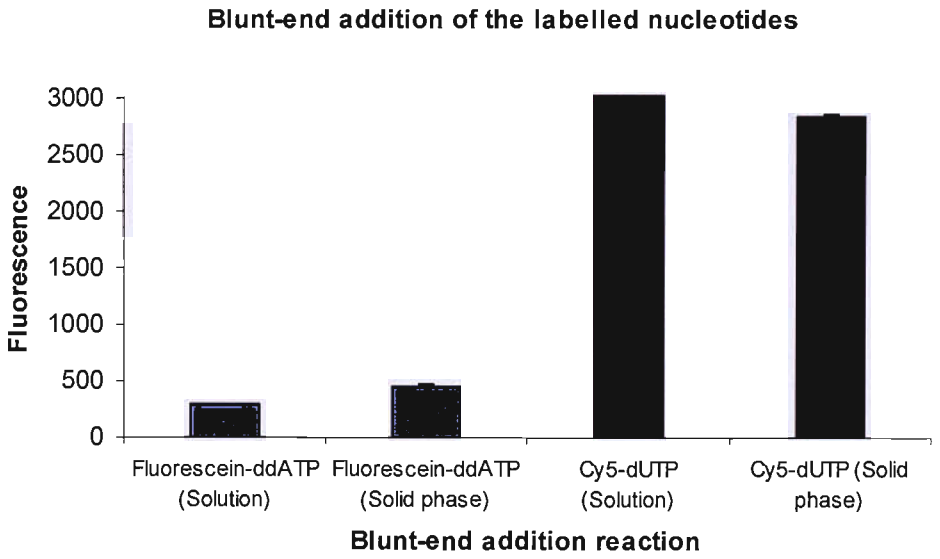


Figure 5.26 Fluorescence (± SE) exhibited by the solution and solid phase blunt-end addition reaction samples.

The low fluorescence results obtained for both solution (290±1.2) and solid phase (454±1.5) blunt-end addition reactions incubated with fluorescein-ddATP, indicated undetectable levels of incorporation. On the other hand, solution (3,006±9.9) and solid phase (2,832±10.5) blunt-end addition reactions incubated with Cy5-dUTP displayed very high fluorescence in both cases indicating the addition of the labelled nucleotide onto the substrate. The controls consisting of avidin coated GMA beads incubated with *Taq* and fluorescein-ddATP or Cy5-dUTP yielded fluorescences of 298±1.1 and 117±0.9 respectively. The solid phase controls containing the immobilised probe incubated with *Taq* and fluorescein-ddATP or Cy5-dUTP yielded fluorescences of 273±1.0 and 120±0.9 respectively. These low fluorescence results obtained for the controls indicate undetectable levels of non-specific binding of the fluorescent nucleotide in both cases and no nucleotide addition to single stranded probe.

The fact that *Taq* fails to add or adds in insignificant amounts fluorescein-ddATP to the blunt-end substrate is not unexpected, taking into account the strong discrimination against fluorescein-ddNTPs incorporation[15] allied to the absence of a template thymidine residue on the DNA target.

The addition of Cy5-dUTP to the blunt-end duplex is also not surprising as *Taq* is widely known for adding dNTPs[35, 74-78] in a non-templated fashion. Amongst the dNTPS, dATP[35, 74-76] is the most efficiently incorporated nucleotide. Since the DNA target did not contain a template overhang, the addition of Cy5-dUTP cannot involve the use of coding information and occurs through a non-templated approach.[35]

5.3.9 Assaying multiplexed incorporation on solid phase

Results gathered so far demonstrate that *Taq* incorporated correctly fluorescein-ddATP onto DNA substrate A (Figure 5.27A) and did not exhibit misincorporation onto DNA substrate B (Figure 5.27B). However, for Cy5-dUTP, both *Taq* and Thermo Sequenase™ revealed incorporation and misincorporation of the labelled nucleotide onto DNA substrates B and A respectively (Figures 5.27B and A).

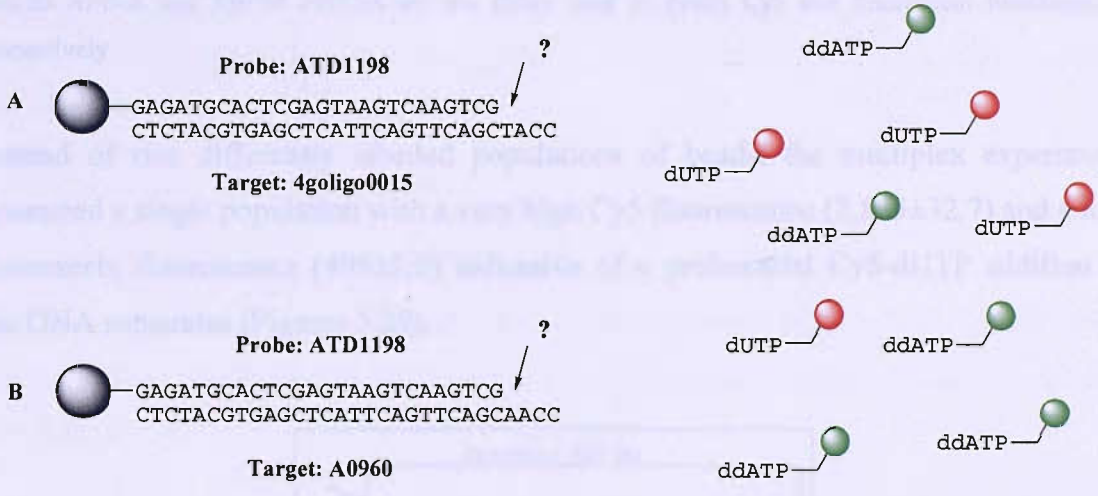


Figure 5.27 Outline for the multiplexed incorporation experiment. For this particular experiment, both DNA substrates A and B underwent pre-hybridisation prior to incubation with enzyme and nucleotides, in order to avoid cross-hybridisation of the targets on the same solid support.

The multiplexed experiment outlined in Figure 5.27, aimed to determine the discrimination ability of *Taq* in the presence of both DNA substrates and both labelled nucleotides. We wanted to investigate if; in the presence of both nucleotides *Taq* would direct them to the correct incorporation sites, generating two differently labelled populations of beads. Figure 5.28 shows the expected result; one population exhibiting high fluorescein fluorescence and low Cy5 fluorescence (P1) indicative of fluorescein-ddATP incorporation and a second population with high Cy5 fluorescence and low fluorescein fluorescence (P2) indicative of Cy5-dUTP incorporation.

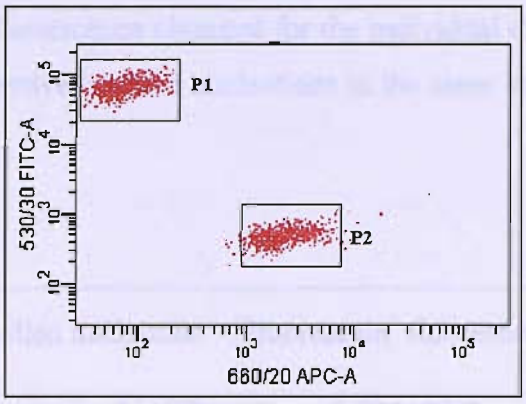


Figure 5.28 Expected FACS dot-plot for the experiment showing two differently labelled populations. 660/20 APC-A and 530/30 FITC-A are the filters used to detect Cy5 and fluorescein fluorescence, respectively.

Instead of two differently labelled populations of beads, the multiplex experiment generated a single population with a very high Cy5 fluorescence (2,555±32.7) and a low fluorescein fluorescence (495±1.9) indicative of a preferential Cy5-dUTP addition to the DNA substrates (Figures 5.29).

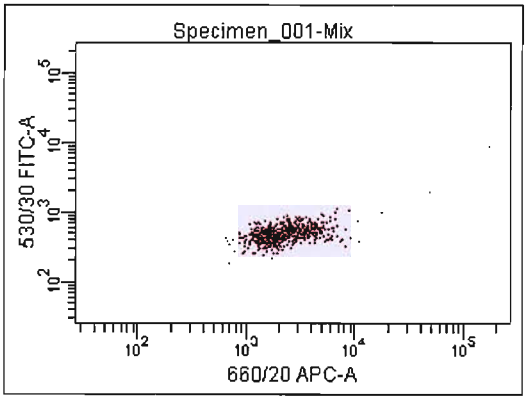


Figure 5.29 FACS dot-plot for the experiment showing one population indicative of a preferential Cy5-dUTP addition to the DNA substrates.

Taq incorporated and misincorporated Cy5-dUTP respectively onto DNA substrates B and A (Figure 5.27) and discriminated against the incorporation of fluorescein-ddATP. A control consisting of both immobilised DNA substrates incubated with both labelled nucleotides yielded a Cy5 fluorescence of 106±0.9 and a fluorescein fluorescence of 262±0.3 indicating no non-specific binding of the nucleotides.

For a comparison, the fluorescence obtained for the individual chain extension reactions incubated with the respective labelled nucleotides in the same conditions is presented in Table 5.1.

Chain extension	Labelled nucleotide	Fluorescein fluorescence	Cy5 fluorescence
Individual chain extension controls	Fluorescein-ddATP	1,004±23.3	111±1.1
	Cy5-dUTP	251±1.3	3,980±56.8

Table 5.1 Fluorescence (± SE) exhibited by the individual chain extension reactions. The DNA duplexes for the individual controls underwent pre-hybridisation prior to incubation with enzyme and respective labelled nucleotide.

To investigate this matter further, DNA substrate with target template specific for the incorporation of fluorescein-ddATP was incubated with both labelled nucleotides (Figure 5.30). We wanted to determine if *Taq* in this competition situation would still discriminate against fluorescein-ddATP incorporation.

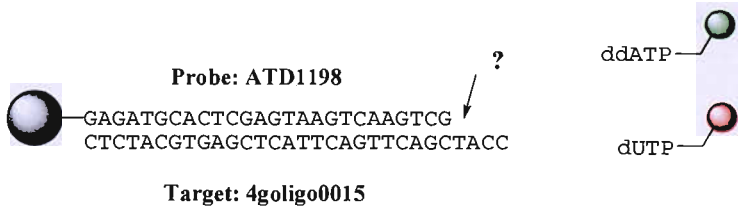


Figure 5.30 Solid phase competition chain extension involving DNA substrate specific for the incorporation of fluorescein-ddATP and both labelled nucleotides.

The competition experiment generated again a single population of beads with a high Cy5 fluorescence (2,097±11.2) and a low fluorescein fluorescence (252±13.9) indicating that *Taq* favoured the misincorporation of Cy5-dUTP.

5.3.10 Time course incorporation of fluorescein-ddATP and Cy5-dUTP in solution and solid phase

The rate of incorporation of fluorescein-ddATP and Cy5-dUTP was compared in solution and on solid phase. The fluorescence exhibited by the beads over the course of the fluorescein-ddATP incorporation experiments is shown in Figure 5.31.

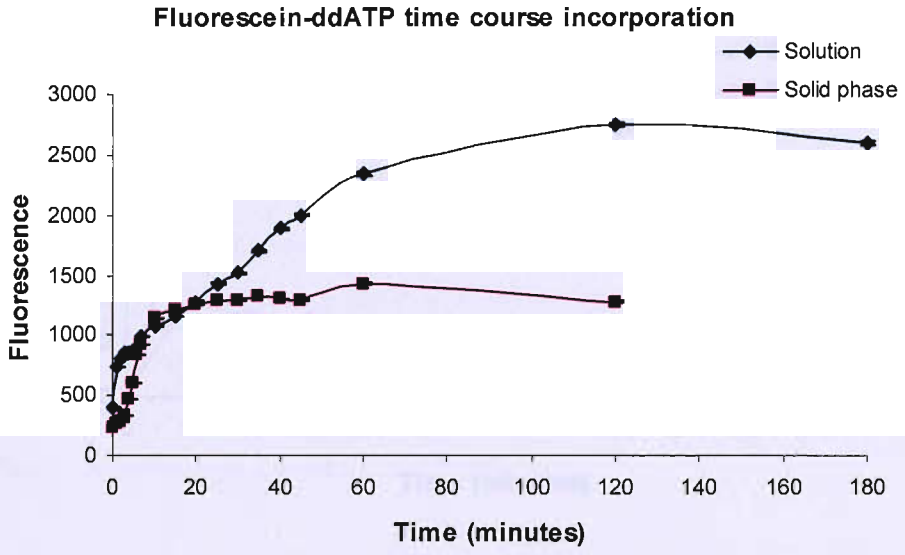


Figure 5.31 Time course experiments in solution and solid phase for the incorporation of fluorescein-ddATP.

The solid phase fluorescence results highlight again the accessibility problem exhibiting a lower yield of nucleotide incorporation. Both incorporation curves seem to go up at the same overall rate, but while the solution experiment shows increasing incorporation of fluorescein-ddATP up to one hour, the solid phase experiment seems to be over after 10 minutes.

The fluorescence shown in Figure 5.32 for the Cy5-dUTP incorporation experiments reveals a much faster incorporation of the nucleotide in solution than on solid phase. Also a greater yield of incorporation is achieved in solution as expected. Both incorporation curves reveal a decreasing trend in fluorescence after reaching its

maximum. This could be interpreted as fluorophore quenching. We have observed that increasing the concentration of Cy5 on GMA beads results in quenching of fluorescence (data not shown). Another possible explanation for the observed decrease in fluorescence could be attributed to a second Cy5-dUTP incorporation. A second Cy5-dUTP incorporation would be expected to occur slowly provoking a quenching of the dye and the observed drop in fluorescence is consistent with this interpretation.

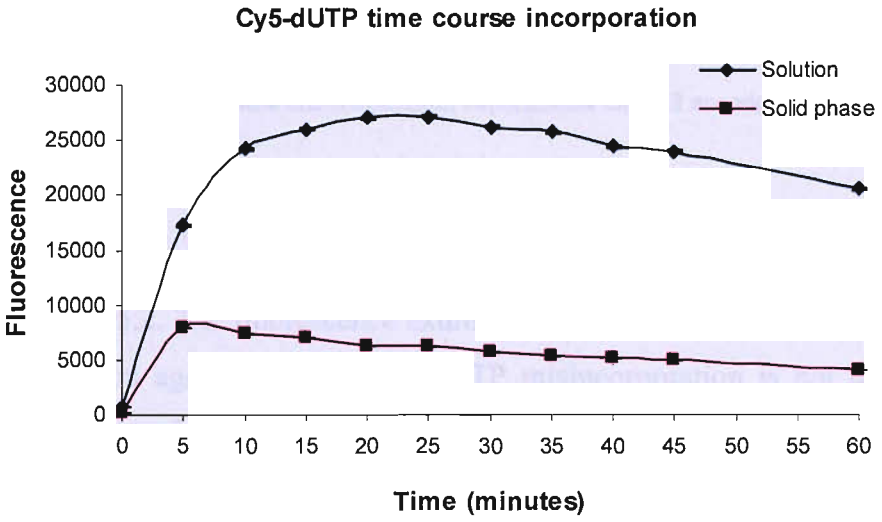


Figure 5.32 Time course experiments in solution and solid phase for the incorporation of Cy5-dUTP.

5.3.11 Single base chain extension on SU8

Single base chain extension reactions using fluorescein-ddNTPs and *Taq*, were attempted on SU8 encoded microparticles as outlined in Figure 5.33. The experiments were carried out in the same manner using the same DNA substrate, differing only in the type of solid support used.

Figure 5.33 Outline of the solid phase chain extension experiments on SU8 encoded microparticles.

An example of two solid phase chain extension experiments carried out on SU8 is displayed in Table 5.2. The fluorescence exhibited by both sets of data is not consistent and a discrimination against fluorescein-ddUTP misincorporation is not observed. In fact, the fluorescence of the samples incubated with fluorescein-ddUTP was higher than those incubated with fluorescein-ddATP in both cases.

Chain Extension	Fluorescein- ddNTP incorporated	Fluorescein fluorescence
Solid phase	ddATP	57,792±801.9
	ddUTP	69,377±980.3
	ddATP	29,023±513.9
	ddUTP	30,225±210.9

Table 5.2 Fluorescence (± SE) of the solid phase chain extension experiments on SU8. Two sets of data from the repetition of the same experiment are presented. A control consisting of avidin coated SU8 barcodes incubated with fluorescein-ddATP displayed a fluorescence of 19,104±292.4.

The fluorescence exhibited by the solution chain extension experiments also did not allow for a distinction between correct versus incorrect nucleotide incorporation as can be seen in Table 5.3.

Chain Extension	Fluorescein- ddNTP incorporated	Fluorescein fluorescence
Solution	ddATP	21,721±396.6
	ddUTP	20,187±375.8

Table 5.3 Fluorescence (± SE) of the preliminary solution chain extension experiments on SU8.

Taq never displayed any evidence of fluorescein-ddNTP misincorporation on avidin coated GMA beads; it is unlikely that the high fluorescence read-out exhibited by SU8 samples could be attributed to that. SU8 barcodes have a high intrinsic fluorescence (13,260±210.9) and a control consisting of avidin coated SU8 barcodes incubated with fluorescein-ddATP displayed an increased fluorescence of 19,104±292.4 indicating a high level of non-specific binding of the nucleotide onto the encoded microparticles. It is likely that this non-specific binding was occurring at different levels generating the high fluorescence observed.

The high standard errors associated with the fluorescence read-out for SU8 encoded particles are due to the lack of homogeneity of the particles as can be seen in Figure 5.34A. A comparison between FACS scatter plots for a population of SU8 and a population of GMA beads used in a chain extension with fluorescein-ddATP is depicted in Figure 5.34.

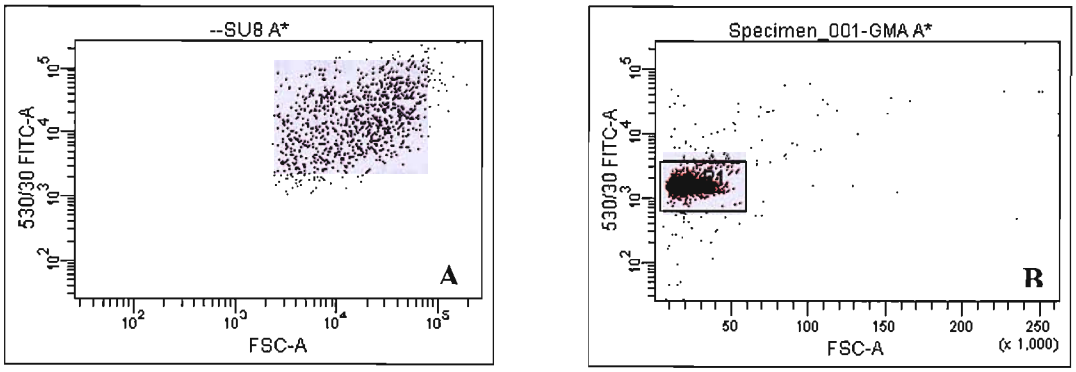


Figure 5.34 A- Scatter plot of a population of SU8 revealing the high scattering of the sample. B- Scatter plot of a population of GMA beads. 530/30 FITC-A is the filter used to detect fluorescein fluorescence and FSC-A (Forward scatter) provides information about the size of the analysed microparticles.

Overall, the fluorescence read-out was inconsistent across experiments and the use of SU8 as solid support was abandoned.

5.3.12 Restriction endonuclease digestion

Previous sections dealt with DNA polymerases and their interaction with different substrates in single base chain extension reactions. This section will focus on a different type of enzymes, restriction endonucleases. Restriction endonucleases as already referred in the introduction of this chapter are enzymes capable of generating DNA fragments and are powerful tools in genetic engineering.[51, 52]

Experiments involving *XhoI*, a restriction endonuclease, were first introduced in Chapter 2. However, the incubation of the enzyme with the tethered DNA construct failed to produce digestion products. This was credited to steric hindrance from the solid support allied to the deficient EDC catalysed attachment of probes which rendered the DNA poorly available for correct hybridisation and inaccessible to enzyme activity.

The introduction of the avidin-biotin system of probe attachment led to a high improvement in hybridisation efficiency and hybridisation discrimination, possibly due to a correct tethering of the DNA probes and a correct orientation for hybridisation. Enzyme accessibility experiments were resumed in this chapter with favourable results thanks to the new probe immobilisation approach.

As for the single base chain extension experiments, the restriction endonuclease digestions described in this section were compared in solution and on solid phase. Figure 5.35 outlines the solid phase restriction experiment. The probe, ATD1198 and the Cy5-labelled target, 4goligo0016 underwent pre-hybridisation prior to immobilisation onto the beads. The beads were then incubated with *XhoI* which promoted the cleavage of the DNA duplex at the restriction site. The endonuclease digestion cleaved off the fluorescent part of the duplex causing the beads to lose fluorescence.

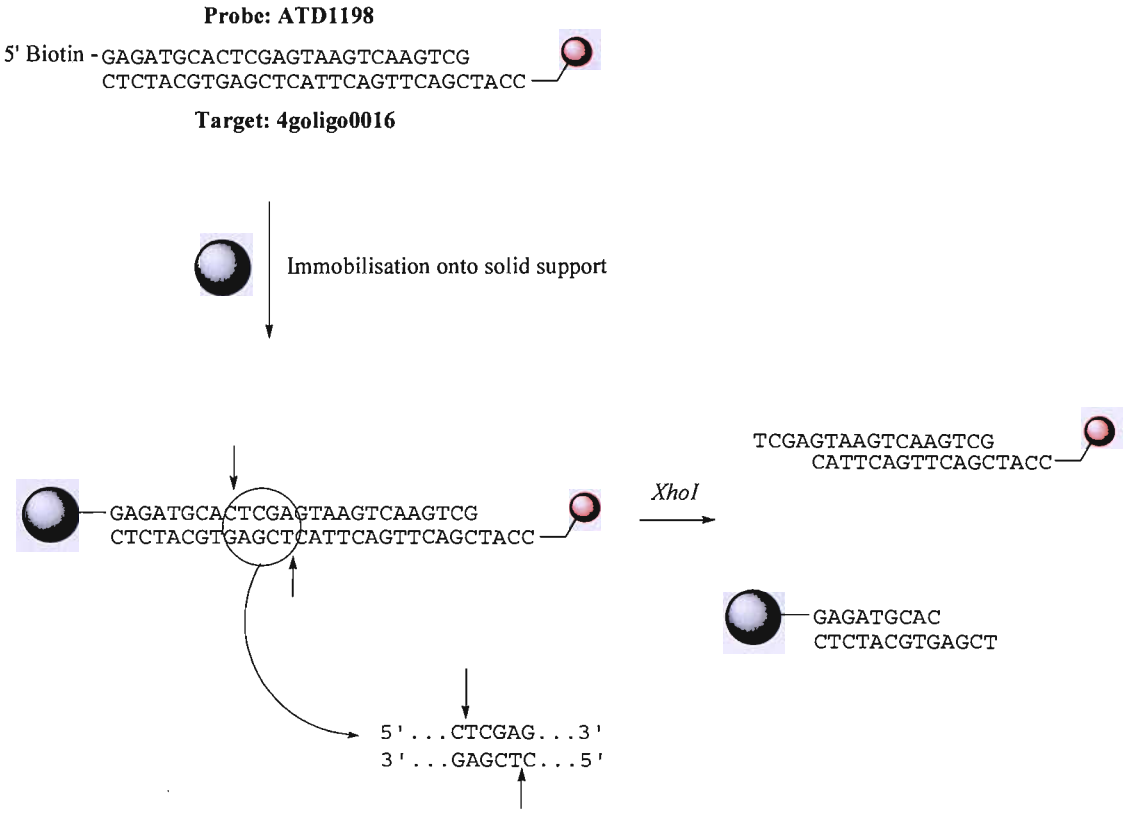


Figure 5.35 Outline of the solid phase restriction endonuclease activity for *XhoI*. The palindromic restriction site is located in the middle of the double helix and is where the enzyme cuts the DNA molecule, between nucleotides C and T in both directions originating two structures with overhangs.

Figure 5.36 outlines the solution approach to the endonuclease cleavage experiment. Both DNA sequences underwent pre-hybridisation prior to incubation with *XhoI* and further incubation with avidin coated GMA beads enabled the immobilisation of the biotinylated fragment rendering non fluorescent beads. After both processes the beads were thoroughly washed to remove any non-specific binding of the labelled fragment and were analysed by FACS.

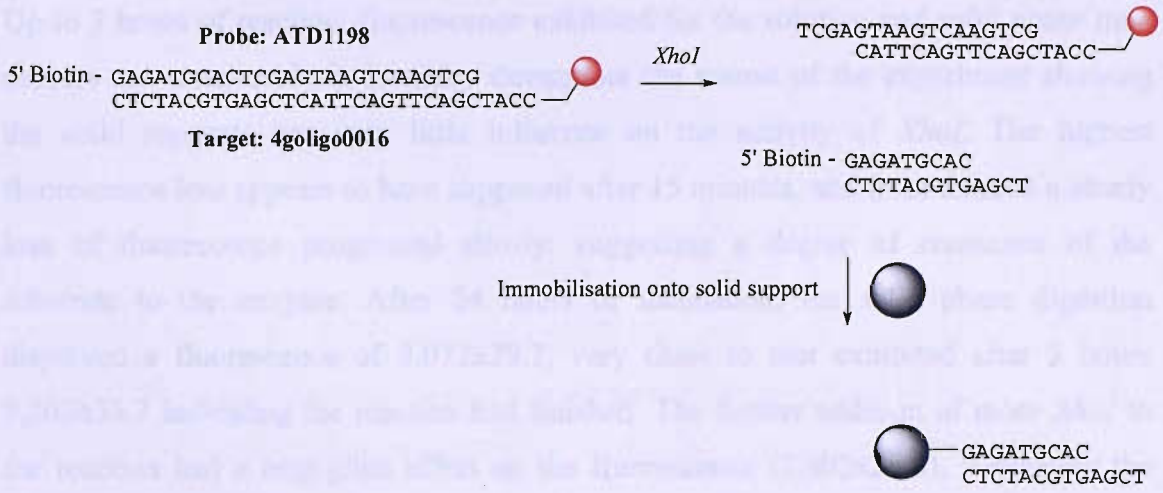


Figure 5.36 Outline of the restriction endonuclease experiment in solution.

The activity of the enzyme in solution and on solid phase was compared. Results are shown in Figure 5.37.

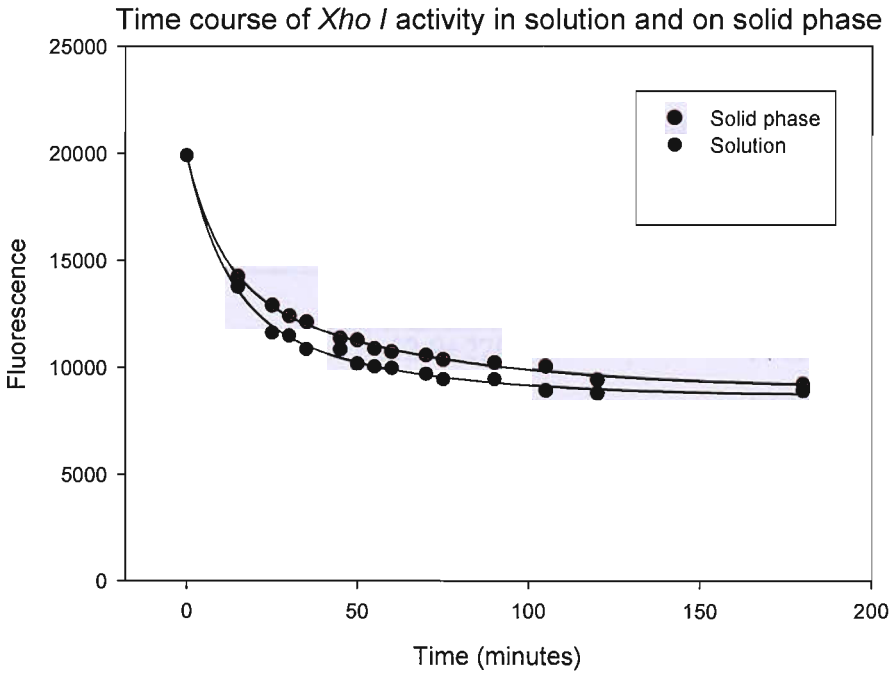


Figure 5.37 Fluorescence of the solution and solid phase endonuclease activity time course experiments. Both data points fit a 5 parameter exponential decay equation $y=A_1e^{-k_1t}+A_2e^{-k_2t}+y_0$.

Up to 3 hours of reaction, fluorescence exhibited for the solution and solid phase time courses are comparable in intensity throughout the course of the experiment showing the solid supports has very little influence on the activity of *XhoI*. The highest fluorescence loss appears to have happened after 15 minutes, and from thereon a steady loss of fluorescence progressed slowly, suggesting a degree of resistance of the substrate to the enzyme. After 24 hours of incubation, the solid phase digestion displayed a fluorescence of 8,072±39.7, very close to that exhibited after 3 hours 9,203±39.3 indicating the reaction had finished. The further addition of more *XhoI* to the reaction had a negligible effect on the fluorescence (7,602±28.6), suggesting the existence of immobilised fluorescent duplexes inaccessible to the enzyme. On the other hand, the solution digestion continued to progress very slowly, exhibiting a slow decrease in fluorescence over the course of 7 days, down to 3,041±17.9 suggesting a slow endonuclease digestion. Similar cases, reporting DNA resistance to *XhoI* cleavage have been described in the literature.[79, 80] A control consisting of a non biotinylated duplex of the same sequence incubated with *XhoI* yielded a fluorescence of 57.2±1.2 indicating the absence of non-specific binding of the labelled fragment.

Figure 5.36 shows that both data points fit a 5 parameter exponential decay equation, $y=A_1e^{-k_1t} + A_2e^{-k_2t} + y_0$. Table 5.4 compares the equation parameters for the exponential decay equation fit in solution and on solid phase. The similarity between the parameters suggests *XhoI* is behaving in the same way both in solution and solid phase.

Equation parameters	Solution	Solid phase
A_1	6652.9±2204.2	5583.9±860.0
A_2	4594.9±1992.4	5333.9±648.8
k_1	0.093±0.03	0.101±0.02
k_2	0.022±0.01	0.017±0.004
y0	8666.0±296.6	8975.9±244.7
R^2	0.9998	0.9999

Table 5.4 Comparison between the equation parameters for the exponential decay equation $y=A_1e^{-k_1t} + A_2e^{-k_2t} + y_0$ fit in solution and on solid phase.

The R^2 value for the 5 parameter exponential decay equation shows a better fit than a 3 parameter exponential decay equation, $y=A_1e^{-k_1t} + y_0$ which generated R^2 values of 0.9992 for the solid phase and 0.9993 for the solution experiments.

Section 5.3.1.3 for the single base chain extension experiments mentioned that the pre-hybridisation of the oligonucleotides before immobilisation onto GMA beads resulted in lower levels of nucleotide incorporation. A similar situation was observed for the digestion experiments. The hybridisation of DNA target to the immobilised probe resulted in higher fluorescence (Figure 5.38) when compared to a pre-hybridisation approach (Figure 5.37). Table 5.5 shows the equation parameters for the exponential decay fit.

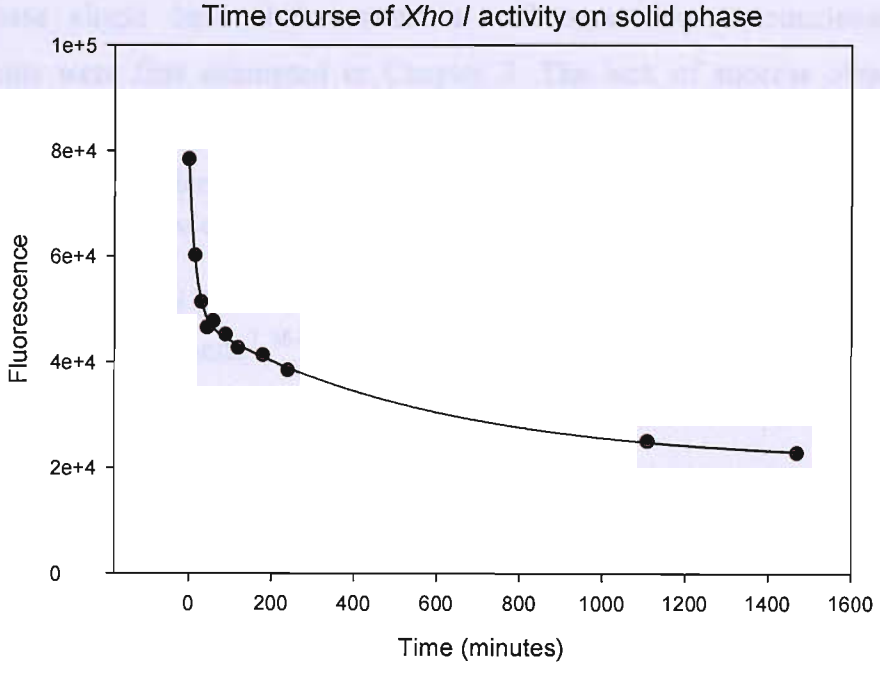


Figure 5.38 Fluorescence of the solid phase time course digestion for the case where the Cy5-labelled target underwent hybridisation to the immobilised probe. Data points fit a 5 parameter exponential decay equation, $y=A_1e^{-k_1t} + A_2e^{-k_2t} + y_0$.

Equation parameters	Solid phase
A_1	30091.7±1622.8
A_2	27463.8±1438.2
k_1	0.06±0.01
k_2	0.002±0.001
y0	20952.6±2097.1
R^2	0.9998

Table 5.5 Parameters for the exponential decay equation $y=A_1e^{-k_1 t} + A_2e^{-k_2 t} + y_0$ fit on solid phase for the case where the Cy5-labelled target underwent hybridisation to the immobilised probe. For comparison, the 3 parameter exponential decay equation generated an R^2 of 0.9864.

5.4 Conclusions

Solid phase single base chain extension and restriction endonuclease digestion experiments were first attempted in Chapter 2. The lack of success obtained in the preliminary experiments was due to the ineffective EDC mediated chemistry of probe attachment which rendered the DNA construct upon hybridisation, poorly available for enzyme accessibility. [63-65] This type of probe attachment was also prone to high levels of non-specific binding of the labelled nucleotides. The introduction of the avidin-biotin system of probe attachment,[7, 66-68] led to an improvement in hybridisation efficiency and hybridisation discrimination. On the basis of that success, these enzyme accessibility experiments were resumed with success.

For the single base chain extension studies two different types of labelled nucleotides were employed to assess and compare the performance of the DNA polymerases both in solution and on solid phase; a fluorescein-ddATP and a Cy5-dUTP. Preliminary experiments in solution revealed that from a set of three DNA polymerases, *Taq*, *Pfu* and *Vent*, *Taq* was the only enzyme able to incorporate a fluorescein-ddATP. These results were not unexpected as *Pfu* and *Vent* exhibit difficulties in incorporating dye-labelled ddNTPs. [27, 28] This is possibly due to their 3'-5' proofreading activity acting in the process.[21, 37] However, the incorporation of fluorescein-ddATP by *Taq* was not very

high. This might be explained as *Taq* is known for strongly discriminating against ddNTPs, especially those labelled with fluorescein.[24-28]

Solid phase chain extensions showed that DNA is accessible for enzyme activity but the level of nucleotide incorporation was lower when compared to that exhibited in solution. This finding is not surprising. Solution base chain extensions are practically universal and require almost no optimisation,[81] however, the transference of this process onto the solid phase introduces problems which are not encountered when the interactions take place in solution as it constraints the way in which the polymerase and substrates interact with each other. DNA is not so free to diffuse as it would be in solution, thus reducing reaction rates as showed by the difference in fluorescence observed.

Taq revealed throughout the course of the experiments to exhibit a high level of selectivity of incorporation of fluorescein-ddATP and a strong discrimination against the misincorporation of fluorescein-ddNTPs both in solution and on solid phase. On the other hand, both *Taq* and Thermo Sequenase™ exhibited high levels of Cy5-dUTP misincorporation in all the experiments performed. In multiplex and competition assays *Taq* favoured the misincorporation of Cy5-dUTP in the presence of fluorescein-ddATP. Non-templated addition of Cy5-dUTP by *Taq* was also verified.[35] This evidence can seriously compromise sequencing methods by generating false positives. Another problem resides in the possibility of multiple nucleotide incorporations directed by the target template (AACC), as suggested by the Cy5-dUTP rate of incorporation experiments both in solution and solid phase.

These experimental results revealed that ddNTPs are more suitable to be used in SNPs identification and DNA analysis methodologies as the level of selectivity of correct nucleotide incorporation was high. DNA polymerases should also be chosen carefully. *Taq* showed to exhibit a low level of fluorescein-ddATP incorporation on solid phase and Thermo Sequenase™ showed to incorporate the labelled nucleotide much more effectively. This agrees with prediction[1, 28] as Thermo Sequenase™ was engineered to improve on ddNTPs incorporation.

Regarding the restriction digestion experiments, the solid support did not seem to have inhibited *XhoI* as the fluorescence exhibited on the solid phase was very close to that in solution. This contrasts with solid phase chain extensions which exhibited lower fluorescence when compared to the solution experiments, clearly suggesting a DNA accessibility issue. This difference in enzyme accessibility on solid phase between *Taq* and *XhoI* can be explained by the fact that the endonuclease is 52 kDa[82] and the DNA polymerase is 94 kDa.[83, 84] This difference in size might explain why *XhoI* could linearly diffuse more easily through the immobilised DNA than *Taq* and reach the restriction recognition site without being hindered by the beads.[85] Conversely, *Taq* being bulkier was sterically hindered by the solid support. The solid phase digestion also suggested the existence of inaccessible immobilised DNA duplexes as the further addition of *XhoI* did not seem to affect the fluorescence of the beads. This evidence might have also affected the performance of the DNA polymerase, contributing to the low fluorescence observed in solid phase base extensions.

The hybridisation approach also seems to play a part in these processes. The pre-hybridisation of the oligonucleotides prior to their immobilisation was showed to result in lower fluorescence. This was observed in both single base extension and restriction digestion experiments. A possible explanation might be a conformational aspect adopted by the duplexes that limits the number of immobilised molecules. The reason why the target hybridisation to the immobilised probe results in higher fluorescence is not understood.

Care should also be taken when choosing the solid support for these reactions as they are known to influence enzyme activity.[86] Solid supports can also display different levels of non-specific binding of fluorescent molecules. Chain extension reactions carried out on SU8 encoded microparticles both in solution and solid phase, failed to provide a significant result as all experiments exhibited high levels of non-specific binding of the fluorophore onto the solid support. Also, the fluorescence exhibited was inconsistent across experiments.

The results presented in this chapter compare the enzymatic performance for single base chain extension and restriction digestion processes in solution and solid phase. Results showed that the nature and modification of the nucleotide can greatly influence the level

of misincorporation compromising sequencing methodologies. The nature and the size of the enzyme were also shown to affect its performance and its accessibility on the immobilised DNA. The type of solid support also revealed to affect the results. All these factors can seriously influence encoded-particle strategies and multiplexing analysis methodologies and must be taken into account if a DNA sequencing platform is to provide reliable results.

5.5 Experimental methods

5.5.1 Reagents

Amino-modified GMA beads (5 - 6 μm - 10% (W/V)) were obtained from Bangs Labs. SU8 encoded microparticles were fabricated in EPFL (Switzerland). Oligonucleotide sequences were synthesised by ATDbio and Sigma-Genosys (Table 5.6). Avidin DN (1.0 mg/ml) was from Vector Laboratories. Fluorescein-12-ddATP / ddCTP / ddGTP / ddUTP were obtained from PerkinElmer Life Sciences. Cy5-dUTP came from Amersham Biosciences. ddTTP came from Fluka. *Taq* DNA polymerase provided with 10x PCR buffer and 25 mM $MgCl_2$ solution, was purchased from Sigma-Aldrich. *Pfu Turbo* DNA polymerase provided with 10x cloned *Pfu* buffer was obtained from Stratagene. *Vent* DNA polymerase provided with 100 mM $MgSO_4$ and 10x ThermoPol Reaction Buffer Pack was obtained from New England Bio Labs Inc. Thermo Sequenase™ DNA polymerase (with *Thermoplasma acidophilum* Inorganic Pyrophosphatase), provided with 1x Thermo Sequenase reaction buffer and 1x Thermo Sequenase dilution buffer was obtained from Amersham Biosciences. *XhoI* provided with 10x NEB 2 buffer came from New England Bio Labs Inc..

DNA	Type of modification	Sequence 5' – 3'	Length (nt)
ATD1198	5' Biotin	GAGATGCACTCGAGTAAGTCAAGTCG	26
A0960	-	CCATCGACTTGACTTACTCGAGTGCATCTC	30
4goligo0015	-	CCATCGACTTGACTTACTCGAGTGCATCTC	30
4goligo0016	5' Cy5	CCATCGACTTGACTTACTCGAGTGCATCTC	30
4goligo26	-	CGACTTGACTTACTCGAGTGCATCTC	26
4goligo0014	5' NH_2	GAGATGCACTCGAGTAAGTCAAGTCG	26

Table 5.6 DNA oligonucleotides employed.

5.5.2 FACS analysis

The experimental methodology for the FACS analysis of the samples of beads was performed as described in section 2.5.2.

5.5.3 Preparation of avidin functionalised solid supports

The experimental methodology for the avidin functionalisation of GMA beads is described in section 2.5.14. Avidin coated SU8 microparticles were obtained from the research group and were prepared as follows: Carboxy modified SU8 microparticles (5 mg) were suspended in EDC coupling solution (600 µl) and agitated at room temperature for 20 minutes. Avidin solution (800 µl, 0.016 µmol) was added and the suspension was agitated for further three hours at 50 °C. The particles were washed and resuspended in 2.5 ml of 5xSSPE, 0.1% SDS buffer.

5.5.5 Single base chain extensions

Solution chain extension incubation reactions in a final volume of 50 µL contained biotinylated probe (ATD1198, 1 µM), target template (4goligo0015 or A0960, 4 µM), labelled nucleotide (fluorescein-ddNTP or Cy5-dUTP, 20 µM), $MgCl_2$ (2.5 mM) and *Taq* polymerase (5 U) in manufacturer's buffer (1x). The mixture was incubated at 60 °C for 1 hour and then 10 µL (33 µg) of avidin coated GMA beads were added and the

suspension was agitated for 30 minutes to promote DNA immobilisation onto the solid supports. The beads were thoroughly washed to remove any non-specific bound nucleotide, resuspended in 5xSSPE, 0.1 % SDS and analysed by FACS.

Solution chain extension incubation reactions with *Vent* (2 U) and *Pfu* (2.5 U) polymerases contained $MgSO_4$ (10 mM) and 5 µL of DMSO respectively, in manufacturer's buffer (1x). The extension reactions were incubated at 60° C and 80° C respectively and in both cases the enzymes were added after heating

Solid phase chain extension incubation reactions involved the pre-immobilisation of biotinylated probe (ATD1198, 1 µM) onto 10 µL (33 µg) of avidin coated GMA beads. The procedure followed thereon was as described for solution incubation reactions. Solid phase extensions involving Thermo Sequenase™ (5 U) were carried out in manufacturer's buffer (0.1x).

Controls consisting of beads or probe immobilised beads incubated with the labelled nucleotides and *Taq* to assess the level of non-specific binding were prepared and submitted to the same experimental reaction conditions.

5.5.6 Competitive chain extension

Solution and solid phase extension reactions were performed as described in section 5.5.5 for *Taq* DNA polymerase and contained fluorescein-ddATP (2 µM) and increasing concentration of dTTP (0, 10, 20, 40, 100 and 200 µM).

Controls consisting of beads or probe immobilised beads incubated with the labelled nucleotide and *Taq* to assess the level of non-specific binding of the fluorescent nucleotide at the working concentration of 2 µM were prepared and submitted to the same experimental reaction conditions. A solution control having a final concentration of 200 µM ddTTP was also prepared to assess the influence of this nucleotide on the fluorescence.

5.5.7 Studying the effect of increasing the concentration of fluorescein-ddATP

Solution and solid phase extension reactions were preformed as described in section 5.5.5 for *Taq* DNA polymerase and contained increasing quantities of fluorescein-ddATP (0.2, 0.4, 2, 2.5, 3.4, 5, 10, 20 μM).

5.5.8 Analysing misincorporation of nucleotides by changing target template

The misincorporation experiments of fluorescein-ddATP onto target A0960 and of Cy5-dUTP onto target 4goligo0015 were carried out as described in section 5.5.5 for *Taq* and Thermo Sequenase™ DNA polymerases by varying the target DNA template utilised in the chain extension incubation.

5.5.9 Non-templated addition of labelled nucleotides

Solution and solid phase non-templated addition reactions were performed as described in section 5.5.5 for *Taq* DNA polymerase and contained target 4goligo26 (4 μM).

5.5.10 Assaying multiplexed incorporation on solid phase

The multiplexed incorporation experiment, involved the pre-hybridisation of targets 4goligo0015 and A0960 to the immobilised probes at room temperature for 15 minutes, prior to incubation with the enzyme and the labelled nucleotides. The beads were then washed to remove non bound DNA and combined in 5 μL of water. The procedure followed thereon is as described in section 5.5.5 for *Taq* DNA polymerase containing equimolar quantities of fluorescein-ddATP and Cy5-dUTP. A control consisting of both immobilised DNA substrates incubated with both labelled nucleotides was prepared in the same way to assess the level of non-specific binding of the nucleotides.

5.5.11 Time course incorporation of Cy5-dUTP and fluorescein-ddATP in solution and solid phase

Solution and solid phase incubation reactions were prepared as described in section 5.5.5 for *Taq* DNA polymerase. Aliquots were taken at designated times to monitor the progress of the incorporation of the nucleotides.

5.5.12 Single base chain extension on SU8

The single base chain extension reactions in solution and solid phases carried out on SU8 with *Taq* DNA polymerase were performed as described for GMA beads in section 5.5.5 and involved 20 µL of the avidin coated SU8 stock suspension. A control consisting of SU8 microparticles and fluorescein-ddATP was prepared and underwent the same experimental reaction conditions.

5.5.13 Restriction endonuclease digestion

A solution containing biotinylated probe DNA (ATD1198, 0.6 µM) and Cy5-labelled target (4goligo0016, 0.6 µM), was prepared in 5xSSPE, 0.1 % SDS, final volume 280 µL. The mixture was heated to 92 °C for 2 minutes and cooled to room temperature to promote duplex hybridisation. It was then divided in 2 samples containing each 140 µL. For the solid phase digestion, one sample (140 µL) was immobilised onto 40 µL (0.13 mg) of avidin coated GMA beads to promote duplex immobilisation. The beads were then washed and resuspended in 140 µL of 5xSSPE, 0.1 % SDS. A 5 µL aliquot was taken and analysed by FACS to assess the fluorescence control for the experiment. To the remaining 135 µL of the suspension, 12 µL of NEB 2 buffer and 10 µL (200 units) of *XhoI* were added. The suspension was incubated at 37 °C with constant agitation. Aliquots were taken at designated times, the beads washed to remove any non-specific binding and analysed by FACS to monitor the progress of the digestion reaction. For the solution phase digestion, the remaining sample of hybridised DNA (140 µL) was incubated with 12 µL of NEB 2 buffer and 10 µL (200 units) of *XhoI*. The suspension was incubated at 37 °C. Aliquots were taken at designated times and immobilised onto avidin coated GMA beads. The beads were washed and analysed by FACS.

A control consisting of pre-hybridised probe 4goligo0014 and target 4goligo0016 incubated with *XhoI* and avidin coated GMA beads for 24 hours, to assess the level of non-specific binding of the labelled fragment, was submitted to the same experimental conditions.

5.6 References

1. Dubiley, S.; Kirillov, E.; Mirzabekov, A., Polymorphism analysis and gene detection by minisequencing on an array of gel-immobilized primers. *Nucleic Acids Res.* **1999**, 27, e19.
2. Pati, N.; Schowinsky, V.; Kokanovic, O.; Magnuson, V.; Ghosh, S., A comparison between SNaPshot, pyrosequencing, and biplex invader SNP genotyping methods: accuracy, cost, and throughput. *J. Biochem. Biophys. Methods* **2004**, 60, 1-12.
3. http://www.pyrosequencing.com/DynPage.aspx?id=8583&mn1=1366&mn2=1370.
4. Shapero, M.H.; Leuther, K.K.; Nguyen, A.; Scott, M.; Jones, K.W., SNP genotyping by multiplexed solid-phase amplification and fluorescent minisequencing. *Genome Res.* **2001**, 11, 1926-34.
5. Curcio, M.; Stalhandske, P.; Lindberg, P.; Roeraade, J., Multiplex high-throughput solid-phase minisequencing by capillary electrophoresis and liquid core waveguide fluorescence detection. *Electrophoresis* **2002**, 23, 1467-72.
6. Pourmand, N.; Elahi, E.; Davis, R.W.; Ronaghi, M., Multiplex Pyrosequencing. *Nucleic Acids Res.* **2002**, 30, e31.
7. Cai, H.; White, P.S.; Torney, D.; Deshpande, A.; Wang, Z.L.; Keller, R.A.; Marrone, B.; Nolan, J.P., Flow cytometry-based minisequencing: A new platform for high- throughput single-nucleotide polymorphism scoring. *Genomics* **2000**, 66, 135-43.
8. Carlson, C.S.; Newman, T.; Nickerson, D.A., SNPing in the human genome. *Curr. Opin. Chem. Biol.* **2001**, 5, 78-85.
9. Tully, G.; Sullivan, K.M.; Nixon, P.; Stones, R.E.; Gill, P., Rapid detection of mitochondrial sequence polymorphisms using multiplex solid-phase fluorescent minisequencing. *Genomics* **1996**, 34, 107-13.
10. Lindroos, K.; Liljedahl, U.; Raitio, M.; Syvanen, A.C., Minisequencing on oligonucleotide microarrays: comparison of immobilisation chemistries. *Nucleic Acids Res.* **2001**, 29, E69-9.
11. Pastinen, T.; Kurg, A.; Metspalu, A.; Peltonen, L.; Syvanen, A.C., Minisequencing: a specific tool for DNA analysis and diagnostics on oligonucleotide arrays. *Genome Res.* **1997**, 7, 606-14.
12. Southern, E.M.; Mir, K.; Shchepinov, M., Molecular interactions on microarrays. *Nat. Genet.* **1999**, 21, 5-9.
13. Nikiforov, T.T.; Rendle, R.; Goelet, P.; Rogers, Y.H.; Kotewicz, M.L.; Anderson, S.; Trainor, G.L.; Knapp, M.R., Genetic Bit Analysis: a solid phase method for typing single nucleotide polymorphisms. *Nucleic Acids Res.* **1994**, 22, 4167-75.
14. Syvanen, A.C.; Aalto-Setala, K.; Harju, L.; Kontula, K.; Soderlund, H., A primer-guided nucleotide incorporation assay in the genotyping of apolipoprotein E. *Genomics* **1990**, 8, 684-92.

15. Brandis, J.W., Dye structure affects Taq DNA polymerase terminator selectivity. *Nucleic Acids Res.* **1999**, 27, 1912-8.
16. Prober, J.M.; Trainor, G.L.; Dam, R.J.; Hobbs, F.W.; Robertson, C.W.; Zagursky, R.J.; Cocuzza, A.J.; Jensen, M.A.; Baumeister, K., A system for rapid DNA sequencing with fluorescent chain-terminating dideoxynucleotides. *Science* **1987**, 238, 336-41.
17. Haff, L.A.; Smirnov, I.P., Single-nucleotide polymorphism identification assays using a thermostable DNA polymerase and delayed extraction MALDI-TOF mass spectrometry. *Genome Res.* **1997**, 7, 378-88.
18. Syvanen, A.C.; Ikonen, E.; Manninen, T.; Bengtstrom, M.; Soderlund, H.; Aula, P.; Peltonen, L., Convenient and quantitative determination of the frequency of a mutant allele using solid-phase minisequencing: application to aspartylglucosaminuria in Finland. *Genomics* **1992**, 12, 590-5.
19. Syvanen, A.C.; Aalto-Setala, K.; Harju, L.; Kontula, K.; Soderlund, H., A primer-guided nucleotide incorporation assay in the genotyping of apolipoprotein E. *Genomics* **1990**, 8, 684-92.
20. Lacenere, C.; Garg, M.K.; Stoltz, B.M.; Quake, S.R., Effects of a modified dye-labeled nucleotide spacer arm on incorporation by thermophilic DNA polymerases. *Nucleosides Nucleotides Nucleic Acids* **2006**, 25, 9-15.
21. Anderson, J.P.; Angerer, B.; Loeb, L.A., Incorporation of reporter-labeled nucleotides by DNA polymerases. *Biotechniques* **2005**, 38, 257-64.
22. Bessman, M.J.; Kornberg, A.; Lehman, I.R.; Simms, E.S., Enzymic synthesis of deoxyribonucleic acid. *Biochim. Biophys. Acta* **1956**, 21, 197-8.
23. Rothwell, P.J.; Waksman, G., Structure and mechanism of DNA polymerases. *Adv. Protein Chem.* **2005**, 71, 401-40.
24. Evans, S.J.; Fogg, M.J.; Mamone, A.; Davis, M.; Pearl, L.H.; Connolly, B.A., Improving dideoxynucleotide-triphosphate utilisation by the hyper-thermophilic DNA polymerase from the archaeon Pyrococcus furiosus. *Nucleic Acids Res.* **2000**, 28, 1059-66.
25. Li, Y.; Mitaxov, V.; Waksman, G., Structure-based design of Taq DNA polymerases with improved properties of dideoxynucleotide incorporation. *Proc. Natl. Acad. Sci. USA* **1999**, 96, 9491-6.
26. Tabor, S.; Richardson, C.C., A single residue in DNA polymerases of the Escherichia coli DNA polymerase I family is critical for distinguishing between deoxy- and dideoxyribonucleotides. *Proc. Natl. Acad. Sci. USA* **1995**, 92, 6339-43.
27. Gardner, A.F.; Jack, W.E., Determinants of nucleotide sugar recognition in an archaeon DNA polymerase. *Nucleic Acids Res.* **1999**, 27, 2545-53.
28. Gardner, A.F.; Jack, W.E., Acyclic and dideoxy terminator preferences denote divergent sugar recognition by archaeon and Taq DNA polymerases. *Nucleic Acids Res.* **2002**, 30, 605-13.
29. Vander Horn, P.B.; Davis, M.C.; Cunniff, J.J.; Ruan, C.; McArdle, B.F.; Samols, S.B.; Szasz, J.; Hu, G.; Hujer, K.M.; Domke, S.T.; Brummet, S.R.; Moffett, R. B.; Fuller, C.W., Thermo Sequenase DNA polymerase and T. acidophilum pyrophosphatase: new thermostable enzymes for DNA sequencing. *Biotechniques* **1997**, 22, 758-62, 764-5.
30. Giller, G.; Tasara, T.; Angerer, B.; Muhlegger, K.; Amacker, M.; Winter, H., Incorporation of reporter molecule-labeled nucleotides by DNA polymerases. I. Chemical synthesis of various reporter group-labeled 2'-deoxyribonucleoside-5'-triphosphates. *Nucleic Acids Res.* **2003**, 31, 2630-5.

31. Tasara, T.; Angerer, B.; Damond, M.; Winter, H.; Dorhofer, S.; Hubscher, U.; Amacker, M., Incorporation of reporter molecule-labeled nucleotides by DNA polymerases. II. High-density labeling of natural DNA. *Nucleic Acids Res.* **2003**, 31, 2636-46.
32. Arezi, B.; Hansen, C.J.; Hogrefe, H.H., Efficient and high fidelity incorporation of dye-terminators by a novel archaeal DNA polymerase mutant. *J. Mol. Biol.* **2002**, 322, 719-29.
33. Li, Y.; Waksman, G., Crystal structures of a ddATP-, ddTTP-, ddCTP, and ddGTP- trapped ternary complex of Klentaq1: insights into nucleotide incorporation and selectivity. *Protein Sci.* **2001**, 10, 1225-33.
34. Patel, P.H.; Kawate, H.; Adman, E.; Ashbach, M.; Loeb, L.A., A single highly mutable catalytic site amino acid is critical for DNA polymerase fidelity. *J. Biol. Chem.* **2001**, 276, 5044-51.
35. Clark, J.M., Novel non-templated nucleotide addition reactions catalyzed by procaryotic and eucaryotic DNA polymerases. *Nucleic Acids Res.* **1988**, 16, 9677-86.
36. Zhong, X.; Patel, S.S.; Werneburg, B.G.; Tsai, M.D., DNA polymerase beta: multiple conformational changes in the mechanism of catalysis. *Biochemistry* **1997**, 36, 11891-900.
37. Perler, F.B.; Kumar, S.; Kong, H., Thermostable DNA polymerases. *Adv. Protein Chem.* **1996**, 48, 377-435.
38. Innis, M.A.; Myambo, K.B.; Gelfand, D.H.; Brow, M.A., DNA sequencing with Thermus aquaticus DNA polymerase and direct sequencing of polymerase chain reaction-amplified DNA. *Proc. Natl. Acad. Sci. USA* **1988**, 85, 9436-40.
39. Tindall, K.R.; Kunkel, T.A., Fidelity of DNA synthesis by the Thermus aquaticus DNA polymerase. *Biochemistry* **1988**, 27, 6008-13.
40. Eckert, K. A.; Kunkel, T.A., High fidelity DNA synthesis by the Thermus aquaticus DNA polymerase. *Nucleic Acids Res.* **1990**, 18, 3739-44.
41. Fakhrai-Rad, H.; Pourmand, N.; Ronaghi, M., Pyrosequencing: an accurate detection platform for single nucleotide polymorphisms. *Hum. Mutat.* **2002**, 19, 479-85.
42. Ronaghi, M.; Karamohamed, S.; Pettersson, B.; Uhlen, M.; Nyren, P., Real-time DNA sequencing using detection of pyrophosphate release. *Anal. Biochem.* **1996**, 242, 84-9.
43. Agah, A.; Aghajan, M.; Mashayekhi, F.; Amini, S.; Davis, R.W.; Plummer, J.D.; Ronaghi, M.; Griffin, P.B., A multi-enzyme model for Pyrosequencing. *Nucleic Acids Res.* **2004**, 32, e166.
44. Baner, J.; Nilsson, M.; Mendel-Hartvig, M.; Landegren, U., Signal amplification of padlock probes by rolling circle replication. *Nucleic Acids Res.* **1998**, 26, 5073-8.
45. Larsson, C.; Koch, J.; Nygren, A.; Janssen, G.; Raap, A.K.; Landegren, U.; Nilsson, M., In situ genotyping individual DNA molecules by target-primed rolling-circle amplification of padlock probes. *Nat. Methods* **2004**, 1, 227-32.
46. Kwiatkowski, M.; Nilsson, M.; Landegren, U., Synthesis of full-length oligonucleotides: cleavage of apurinic molecules on a novel support. *Nucleic Acids Res.* **1996**, 24, 4632-8.
47. Szemes, M.; Bonants, P.; de Weerdt, M.; Baner, J.; Landegren, U.; Schoen, C.D., Diagnostic application of padlock probes--multiplex detection of plant pathogens using universal microarrays. *Nucleic Acids Res.* **2005**, 33, e70.
48. Lizardi, P.M.; Huang, X.; Zhu, Z.; Bray-Ward, P.; Thomas, D.C.; Ward, D.C., Mutation detection and single-molecule counting using isothermal rolling-circle amplification. *Nat. Genet.* **1998**, 19, 225-32.

49. Zhang, D.; Wu, J.; Ye, F.; Feng, T.; Lee, I.; Yin, B., Amplification of circularizable probes for the detection of target nucleic acids and proteins. *Clin. Chim. Acta* **2006**, 363, 61-70.
50. Baner, J.; Isaksson, A.; Waldenstrom, E.; Jarvius, J.; Landegren, U.; Nilsson, M., Parallel gene analysis with allele-specific padlock probes and tag microarrays. *Nucleic Acids Res.* **2003**, 31, e103.
51. Che, Y.; Chen, X., A multiplexing single nucleotide polymorphism typing method based on restriction-enzyme-mediated single-base extension and capillary electrophoresis. *Anal. Biochem.* **2004**, 329, 220-9.
52. Zheleznaya, L.; Shiryaev, S.; Zheleznyakova, E.; Matvienko, N.; Matvienko, N., R.SspD5I is a neoschizomer of HphI producing blunt end DNA fragments. *FEBS Lett.* **1999**, 448, 38-40.
53. Cooper, G.M., The Cell: A Molecular Approach. **2000**, Sinauer Associates, Inc., Sunderland, Missachusetts ISBN 0-87893-106-62000, (2nd edn.).
54. Hamilton, M.D.; Evans, D.H., Enzymatic processing of replication and recombination intermediates by the vaccinia virus DNA polymerase. *Nucleic Acids Res.* **2005**, 33, 2259-68.
55. Reinartz, J.; Bruyns, E.; Lin, J. Z.; Burcham, T.; Brenner, S.; Bowen, B.; Kramer, M.; Woychik, R., Massively parallel signature sequencing (MPSS) as a tool for in-depth quantitative gene expression profiling in all organisms. *Brief. Funct. Genomic. Proteomic.* **2002**, 1, 95-104.
56. Brenner, S.; Johnson, M.; Bridgham, J.; Golda, G.; Lloyd, D.H.; Johnson, D.; Luo, S.J.; McCurdy, S.; Foy, M.; Ewan, M.; Roth, R.; George, D.; Eletr, S.; Albrecht, G.; Vermaas, E.; Williams, S.R.; Moon, K.; Burcham, T.; Pallas, M.; DuBridge, R.B.; Kirchner, J.; Fearon, K.; Mao, J.; Corcoran, K., Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. *Nat. Biotechnol.* **2000**, 18, 630-4.
57. Brenner, S.; Williams, S.R.; Vermaas, E.H.; Storck, T.; Moon, K.; McCollum, C.; Mao, J.I.; Luo, S.J.; Kirchner, J.J.; Eletr, S.; DuBridge, R.B.; Burcham, T.; Albrecht, G., In vitro cloning of complex mixtures of DNA on microbeads: Physical separation of differentially expressed cDNAs. *Proc. Natl. Acad. Sci. USA* **2000**, 97, 1665-70.
58. Zhang, J.; Li, K.; Liao, D.; Pardinas, J.R.; Chen, L.; Zhang, X., Different applications of polymerases with and without proofreading activity in single-nucleotide polymorphism analysis. *Lab. Invest.* **2003**, 83, 1147-54.
59. Liu, W.H.; Kaur, M.; Makrigiorgos, G.M., Detection of hotspot mutations and polymorphisms using an enhanced PCR-RFLP approach. *Hum. Mutat.* **2003**, 21, 535-41.
60. Szantai, E.; Szmola, R.; Sasvari-Szekely, M.; Guttman, A.; Ronai, Z., The polymorphic nature of the human dopamine D4 receptor gene: a comparative analysis of known variants and a novel 27 bp deletion in the promoter region. *BMC Genet.* **2005**, 6, 39-51.
61. Thiel, T.; Kota, R.; Grosse, I.; Stein, N.; Graner, A., SNP2CAPS: a SNP and INDEL analysis tool for CAPS marker development. *Nucleic Acids Res.* **2004**, 32, e5.
62. Zhang, R.; Zhu, Z.; Zhu, H.; Nguyen, T.; Yao, F.; Xia, K.; Liang, D.; Liu, C., SNP Cutter: a comprehensive tool for SNP PCR-RFLP assay design. *Nucleic Acids Res.* **2005**, 33, (Web Server issue), W489-92.
63. Adessi, C.; Matton, G.; Ayala, G.; Turcatti, G.; Mermod, J.J.; Mayer, P.; Kawashima, E., Solid phase DNA amplification: characterisation of primer attachment and amplification mechanisms. *Nucleic Acids Res.* **2000**, 28, e87.

64. Laboratories, B., TechNote 205 Rev. 003, Active: 30/MAR/02 - Covalent Coupling, Published by Bangs Laboratories, Inc.
65. Walsh, M.K.; Wang, X.W.; Weimer, B.C., Optimizing the immobilization of single-stranded DNA onto glass beads. *J. Biochem. Biophys. Methods* **2001**, 47, 221-231.
66. Ali, M.F.; Kirby, R.; Goodey, A.P.; Rodriguez, M.D.; Ellington, A.D.; Neikirk, D.P.; McDevitt, J.T., DNA hybridization and discrimination of single-nucleotide mismatches using chip-based microbead arrays. *Anal. Chem.* **2003**, 75, 4732-39.
67. Goto, S.; Takahashi, A.; Kamisango, K.; Matsubara, K., Single-nucleotide polymorphism analysis by hybridization protection assay on solid support. *Anal. Biochem.* **2002**, 307, 25-32.
68. Pirrung, M.C.; Worden, J.D.; Labriola, J.P.; Montague-Smith, M.P.; Weislo, L.J., Solid-phase, single nucleotide primer extension of DNA/RNA hybrids by reverse transcriptases. *Bioorg. Med. Chem. Lett.* **2001**, 11, 2437-40.
69. McQuain, M.K.; Seale, K.; Peek, J.; Fisher, T.S.; Levy, S.; Stremler, M.A.; Haselton, F.R., Chaotic mixer improves microarray hybridization. *Anal. Biochem.* **2004**, 15, 215-26.
70. Schaupp, C.J.; Jiang, G.; Myers, T.G.; Wilson, M.A., Active mixing during hybridization improves the accuracy and reproducibility of microarray results. *Biotechniques* **2005**, 38, 117-9.
71. Murphy, D., Gene expression studies using microarrays: principles, problems, and prospects. *Adv. Physiol. Educ.* **2002**, 26, 256-70.
72. Mujumdar, R.B.; Ernst, L.A.; Mujumdar, S.R.; Lewis, C.J.; Waggoner, A.S., Cyanine dye labeling reagents: sulfoindocyanine succinimidyl esters. *Bioconjug. Chem.* **1993**, 4, 105-11.
73. Yu, H.; Chao, J.; Patek, D.; Mujumdar, R.; Mujumdar, S.; Waggoner, A. S., Cyanine dye dUTP analogs for enzymatic labeling of DNA probes. *Nucleic Acids Res.* **1994**, 22, 3226-32.
74. Brownstein, M.J.; Carpten, J.D.; Smith, J.R., Modulation of non-templated nucleotide addition by Taq DNA polymerase: primer modifications that facilitate genotyping. *Biotechniques* **1996**, 20, 1004-6, 1008-10.
75. Culjkovic, B.; Savic, D.; Stojkovic, O.; Romac, S., Poly(A) tailing of ancient DNA: a method for reproducible microsatellite genotyping. *Anal. Biochem.* **2003**, 318, 124-31.
76. Magnuson, V.L.; Ally, D.S.; Nylund, S.J.; Karanjawala, Z.E.; Rayman, J.B.; Knapp, J.I.; Lowe, A.L.; Ghosh, S.; Collins, F.S., Substrate nucleotide-determined non-templated addition of adenine by Taq DNA polymerase: implications for PCR-based genotyping and cloning. *Biotechniques* **1996**, 21, 700-9.
77. Smith, J.R.; Carpten, J.D.; Brownstein, M.J.; Ghosh, S.; Magnuson, V.L.; Gilbert, D.A.; Trent, J.M.; Collins, F.S., Approach to genotyping errors caused by nontemplated nucleotide addition by Taq DNA polymerase. *Genome Res.* **1995**, 5, 312-7.
78. Yang, S.W.; Astatke, M.; Potter, J.; Chatterjee, D.K., Mutant Thermotoga neapolitana DNA polymerase I: altered catalytic properties for non-templated nucleotide addition and incorporation of correct nucleotides. *Nucleic Acids Res.* **2002**, 30, 4314-20.
79. Godwin, A.R.; Bollag, R.J.; Christie, D.M.; Liskay, R.M., Spontaneous and restriction enzyme-induced chromosomal recombination in mammalian cells. *Proc. Natl. Acad. Sci. USA* **1994**, 91, 12554-8.

80. Kwoh, T.J.; Kwoh, D.Y.; McCue, A.W.; Davis, G.R.; Patrick, D.; Gingeras, T.R., Introduction and expression of the bacterial PaeR7 methylase gene in mammalian cells. *Proc. Natl. Acad. Sci.USA* **1986**, 83, 7713-7.
81. Chen, X.N.; Levine, L.; Kwok, P.Y., Fluorescence polarization in homogeneous nucleic acid analysis. *Genome Res.* **1999**, 9, (5), 492-498.
82. http://www.neb.uk.com/productcatalogue/productinfo.aspx?id=faq-R0146@UK@GUEST@1@@XX@#1100.
83. Lawyer, F.C.; Stoffel, S.; Saiki, R.K.; Chang, S.Y.; Landre, P.A.; Abramson, R.D.; Gelfand, D.H., High-level expression, purification, and enzymatic characterization of full-length Thermus aquaticus DNA polymerase and a truncated form deficient in 5' to 3' exonuclease activity. *PCR Methods Appl.* **1993**, 2, 275-87.
84. Lawyer, F.C.; Stoffel, S.; Saiki, R.K.; Myambo, K.; Drummond, R.; Gelfand, D.H., Isolation, characterization, and expression in Escherichia coli of the DNA polymerase gene from Thermus aquaticus. *J. Biol. Chem.* **1989**, 264, 6427-37.
85. Jeltsch, A.; Wenz, C.; Stahl, F.; Pingoud, A., Linear diffusion of the restriction endonuclease EcoRV on DNA is essential for the in vivo function of the enzyme. *EMBO J.* **1996**, 15, 5104-11.
86. Kress, J.; Zanaletti, R.; Amour, A.; Ladlow, M.; Frey, J.G.; Bradley, M., Enzyme accessibility and solid supports: which molecular weight enzymes can be used on solid supports? An investigation using confocal Raman microscopy. *Chemistry* **2002**, 8, 3769-72.

6. Conclusions

6.1 Introduction

Biological and medical researches have evolved significantly over recent years, with the large-scale screening of whole genomes complementing focused studies on a few genes or proteins. This evolution has included applications ranging from functional analysis of unknown genes to identification of disease-related genes, screening in drug discovery and clinical diagnostics. There has been a parallel surge in technology development to facilitate large-scale biological analysis. Perhaps the most powerful and versatile tool to addressing and separately quantifying the amount of genetic information available in this "post-genomic" era has been provided by planar DNA microarrays. Although planar DNA arrays have had a major impact on high-density high-throughput screening, the quality of the results and the speed at which they can be obtained are limited by the intrinsic properties of the planar features. In recent years, suspension arrays of encoded microparticles have emerged as a promising and an alternative tool to overcome the drawbacks imposed by the planar platform (Chapter 1).

6.2 Summary

The work reported here is part of the basic technology grant (4G project) to design and implement a new method of synthesising, screening and sequencing DNA. [1] The 4G project consists of a suspension array platform involving the creation of libraries of oligonucleotide probes immobilised on the surface of three-dimensional encoded solid supports. These tethered oligonucleotide probes will be used to screen complex genomic pools of DNA targets and will hybridise to their complementary sequences if they are present. The detection of complementary hybridisation will be achieved by single fluorescent nucleotide extension of the immobilised duplex using a DNA

polymerase. The labelled solid supports will be sorted and decoded identifying the sequences of the attached probes and targets. Sequence reassembly will be carried out by computational strategies. It is therefore of paramount importance to study and understand oligonucleotide hybridisation selectivity. This means the ability of one specific DNA probe to recognise and hybridise to its complementary target in a complex mixture of DNA and discriminate against non-complementary targets. Also, it is of interest to study the enzymatic interaction between DNA polymerases and DNA duplexes which are behind the selective incorporation of a nucleotide into double stranded DNA.

The contribution of this thesis to the 4G project involved the exploration of an appropriate chemistry of attachment of DNA probes to solid supports (Chapter 2), optimisation of hybridisation selectivity (Chapters 3 and 4) and single fluorescent nucleotide addition to the immobilised duplex by a DNA polymerase (Chapter 5).

6.2.1 Chemistry of probe attachment; Influence on target hybridisation and enzyme accessibility

According to Wilson *et al.*[2] the amount of molecular information that can be derived from suspension arrays of encoded microparticles is limited by the number of codes that can be distinguished in the same sample. Although this argument is true, many more factors exist that can lead to a limit in molecular information. The work described in this thesis highlights some of those factors.

The first factor that should be considered is the solid support chosen for the immobilisation of probes. The properties of the solid support, such as the constituent material, electrical properties, surface functionalisation, etc., will influence the nature of the selective binding interaction. This was seen in chapter 2 and 5. Amino functionalised silica and GMA beads and avidin functionalised encoded SU8 microparticles exhibited a strong non-specific binding of fluorescein-ddATP. Also, carboxy functionalised un-encoded SU8 microparticles exhibited a strong non-specific binding of fluorescent DNA target. On the other hand, carboxy and avidin

functionalised GMA beads showed very low non-specific binding of labelled nucleotides and fluorescent DNA targets.

The array quality, whether it is a planar or a suspension array, depends on the probe array quality which in turn is intimately connected to the probe attachment chemistry employed. The probe attachment chemistry can have serious implications on the extent of target hybridisation and enzyme accessibility. This was seen in chapter 2. The carbodiimide mediated amide coupling used to tether the amino termini of the oligonucleotides onto carboxy functionalised GMA beads resulted in a high yield of immobilised probe (98%), however, a comparatively low yield of target hybridisation (7%) was observed. Unsuccessful single base extension and restriction digestion of the immobilised constructs suggested steric hindrance from the solid support not allowing the enzymes to access the duplexes. These findings prompted the use of spacers, structures inserted between the solid support and the probes enabling a solution-like environment for the probes and easy access for the targets. However, the results obtained did not show an improvement on hybridisation signal or enzymatic activity. An investigation into the carbodiimide chemistry of probe attachment revealed that more than half of the probes were being tethered through the side chain moieties. This incorrect probe orientation resulted in limited target hybridisation and poor enzyme accessibility. On the other hand the avidin-biotin system of probe attachment resulted in an improvement in hybridisation efficiency and enzyme accessibility. This improvement allowed the studies carried out in subsequent chapters.

6.2.2 Post-hybridisation stringency

Post-hybridisation washes are rarely given the importance they deserve in the literature and they were seen here to play a significant effect in hybridisation discrimination. The post-hybridisation washing stringency protocol developed in this work consisted in washing the beads with buffers and urea solutions of increased stringency. Results showed that for the short probes investigated, washing the beads with buffers of increased stringency was sufficient to attain hybridisation discrimination against completely unrelated targets.

However, in order to discriminate against closely related sequences, urea solutions were also employed. The results shown in chapter 4 for the fluorescence read-out of the comprehensive set of mismatches with increasing washing stringency revealed that hybridisation selectivity was best after 2 M urea. Some exceptions were found in chapter 3 (16-mers) and 4 (double mismatch) in which the increase in washing stringency did not improve discrimination and promoted severely duplex denaturation. This was surprising and suggested a complex interaction between the sequence mutation and the washings.

6.2.3 Hybridisation selectivity

The sequence and length of the probes is of paramount importance in the performance of an array. However, choosing optimal probes has been difficult because of the present poor understanding of how sensitivity and specificity of a probe depend upon its length and sequence. Also, the process of hybridisation still involves a myriad of variables not yet predictable by computational methods and it has been difficult to transfer these findings and predictions to arrays.[3]

The idea that short oligonucleotides with a mismatch should hybridise less efficiently than a perfect match in the context of arrays is not always verified.[4] The results presented in chapter 4 show that the stability exhibited by a mismatch in an array context may differ from that predicted by solution-based hybridisations. The thermodynamic stability predicted by the T_m for a given mismatched base pair may not be reflected on the kinetics of dissociation. For example, the C:C mismatch showed the lowest solution T_m determined of all the mismatches. However, in the context of the array, C:C showed to be quite stable. According to the results in chapter 4, T_m was considered to be a poor predictor of fluorescence hybridisation signal. This is in agreement with the findings of Pozhitkov *et al.*[5] and according to the authors, thermodynamic criteria should not be taken into account when designing oligonucleotide probes.

Bhattacharya and co-workers[6] investigated the kinetics of base pair opening of centrally located homologous mismatches in 9-mer duplexes by ^{1}H NMR studies of imino proton exchange rates. The imino protons of the aromatic heterocyclic base exchange with the solvent protons when the hydrogen bond in the base pair is disrupted. The kinetics of base pair opening trend found by the authors did not correlate the thermodynamic stability trend predicted by the T_m. Both trends were in very good agreement with the fluorescence read-out and T_m results presented in chapter 4 for the homologous mismatches studied. This reinforced the suggestion that the distinction achieved between all matches and all mismatches was due to the higher kinetics of dissociation of the mismatches under the washing conditions employed, rather than discrimination based in thermodynamic stability. The combination of the results shown in chapter 4 with the findings of Bhattacharya and co-workers suggests that shorter base pair lifetimes result in an increased disruption of the π stacked array which translates into faster kinetics of base pair opening and interaction with urea. This disruption also affects the neighbouring bases and ultimately leads to higher kinetics of duplex dissociation. Future work can involve the measurement of the loss of fluorescence read-out of each mismatched base pair with time in urea solutions. It is anticipated that the most kinetically labile base pairs will dissociate more rapidly in these conditions. Also, a similar ^{1}H NMR study for all the 20-mer duplexes used in chapter 4 can be undertaken in order to investigate a correlation with the fluorescence read-out observed.

The type of mismatch, location and context dependence on hybridisation signal are a long way from being understood. This was seen in chapters 3 (16-mers) and 4 (double mismatch), where certain mismatches behaved opposite to what was anticipated. This makes difficult to predict the hybridisation performance of oligonucleotide probes and has serious implications on probe design and consequently has serious implications on accurate expression measurements.

The multiplex assays studied in chapter 3 involving completely unrelated probes and one or two targets exhibited different populations consistent with the individual hybridisation reactions. However, an increase in probe relatedness leads to a decrease in hybridisation discrimination. The multiplexes involving the SNP containing 14-mers exhibited poor hybridisation discrimination and poor population discrimination consistent with individual assays. Also, one of the multiplex assays investigated in

chapter 4 raised the concern that in a SNP discrimination assay, the formation of a stable mismatch may mask the presence of a particular allele. 20-mer probe (T) in the presence of two targets (containing G and A, respectively) favoured the T:G mismatch over the T:A match. The multiplex result did not correlate the selectivity of hybridisation observed in the individual reactions. These situations can compromise an array platform as they can lead to inaccurate gene expression profiling.

6.2.4 Comparison between enzyme accessibility in solution and on solid phase

DNA polymerases and restriction endonucleases are powerful tools used in a multitude of solution and solid phase methodologies for DNA analysis (Chapters 1 and 5). We thought it would be worth to investigate and compare the performance of these enzymes in single base chain extension and restriction digestion processes in solution and solid phase.

The experimental results obtained for single base chain extension reactions showed that the nature and modification of the nucleotide can severely influence the level of misincorporation compromising sequencing methodologies. Results revealed that ddNTPs are more adequate than dNTPs to be used in SNPs identification and DNA analysis methodologies as seen by the high level of selectivity of correct nucleotide incorporation. DNA polymerases should also be chosen carefully. Results showed that Thermo Sequenase™ incorporated fluorescein-ddATP on the solid phase much more effectively than *Taq*.

The size of the enzyme also showed to affect its performance and its accessibility on the immobilised DNA. In restriction digestion experiments, the solid phase fluorescence results were very similar to that exhibited in solution suggesting the solid support did not inhibit *XhoI*. This contrasts with solid phase chain extensions which exhibited lower fluorescence when compared to the solution experiments, clearly suggesting a DNA accessibility issue. This difference in solid phase accessibility for both enzymes can be

explained by a difference in size. The restriction enzyme is a much smaller enzyme than the DNA polymerase.

6.3 Final conclusions

The work described here presents two different approaches to interrogating DNA on the solid phase in the form of a suspension array. The first approach involved direct hybridisation of immobilised probes to targets in solution and relies on differential hybridisation to attain complementary selectivity. The second approach involved the interrogation of the tethered DNA, either by a DNA polymerase or by an endonuclease as a means to evaluate the enzymatic accessibility of the immobilised duplexes. The summary of this work poses two important questions. Are these two approaches viable for analysing DNA? Are suspension arrays suitable to undertake these analyses?

The answer to both questions is obviously yes, otherwise researchers and scientists would have stopped using suspension arrays by now. In fact, quite the contrary has happened. Suspension arrays are likely to become one of the mainstays of future research and clinical applications as they combine cost effective high-throughput genotype screening and accuracy and allow assay platform miniaturisation, a high level of multiplexing and flexible automation. Also, the introduction of encoded solid supports will provide important additional information about the systems studied. However, as with all the technologies developed for DNA analysis, much care and attention need to be directed to the fact that potential interfering factors exist and might compromise the detection signal. The work described here shows some of those potential interfering factors and how they can hinder accurate gene expression measurements.

The presence of a three-dimensional solid support introduces variables that interfere with DNA analysis. Moreover, DNA on the solid phase shows a different behaviour to that exhibited in solution. This was seen here. The stability trend exhibited by the DNA duplexes studied in solution was not replicated on the solid phase. Consequently, solution hybridisation studies fail to predict and understand this solid phase behaviour.

This can lead to unexpected and misleading results which can compromise SNP genotyping assays. The presence of a particular allele may be masked from detection. This has serious implications when we think about multiplex assays involving millions of immobilised probes interrogating millions of targets. Examples of these problems were seen in this work. There were clearly specific cases where we got the wrong answers. All this evidence put together is quite worrying in that, how can probes be correctly designed if do not behave as expected and predicted, when they are immobilised? Unfortunately, the answer to this question is still far from being known.

The enzymatic accessibility study also revealed potential concerning issues. All the experiments carried out using a Cy5-dUTP showed a high level of nucleotide misincorporation, both in solution and solid phase. This dNTP misincorporation was carried out by both the very cheap *Taq* and by the very expensive, high-fidelity polymerase, Thermo Sequenase™. This raises concern, if we think that when DNA target is being labelled and amplified by PCR, the wrong nucleotide incorporation may be propagated throughout the amplification cycles. This will therefore lead to the creation of new interfering target sequences and will definitely compromise the DNA detection platform. Also, the size of the enzyme is an issue to the accessibility of the tethered DNA constructs. However, the differences obtained by the single base chain extension and the endonuclease digestion of the immobilised duplexes in solution and on solid phase may be explained by enzymology and sterics of the processes.

These are some of the most pressing interfering factors limiting the potential of the array technology. One lesson to be derived is that so much still needs to be understood about the processes of DNA hybridisation and all the variables affecting it on the solid phase. Also, the interaction between DNA polymerases and substrates needs to be looked into to fully understand the reasons that lead to nucleotide misincorporation. Comprehensive studies addressing these unknown variables are critical to understanding the function and regulation of all genes and proteins in order to decipher the underlying workings of the cell, determine disease mechanisms and optimise drug targeting.[7]

A few years ago, the array technology was a promise of the future. That future has arrived and already new goals and objectives lie ahead of us. In the same way, it is likely that the future will bring the answers to all the questions raised here and

resolutions to all these interfering factors and more. Mastering that knowledge will ensure the ability to accurately design libraries of probes and to predict and calculate cross-hybridisation. This will help the success of sequencing projects such like the 4G. At the current speed of technology and science development, probably in a near future we will be able to provide the means to a practical individual genome sequencing and the means to a personalised medicine on the basis of genetic information.

6.4 References

1. Bradley M. et al., Basic Technology grant proposal - 4 billion bases a day - practical individual genome sequencing. **2002**.
2. Wilson, R.; Cossins, A.R.; Spiller, D.G., Encoded microcarriers for high-throughput multiplexed detection. *Angew. Chem. Int. Ed. Engl.* **2006**, 45, 6104-17.
3. Chou, C.C.; Chen, C.H.; Lee, T.T.; Peck, K., Optimization of probe length and the number of probes per gene for optimal microarray analysis of gene expression. *Nucleic Acids Res.* **2004**, 32, e99.
4. Naef, F.; Magnasco, M.O., Solving the riddle of the bright mismatches: labeling and effective binding in oligonucleotide arrays. *Phys. Rev. E. Stat. Nonlin. Soft Matter Phys.* **2003**, 68, 011906.
5. Pozhitkov, A.; Noble, P.A.; Domazet-Loso, T.; Nolte, A.W.; Sonnenberg, R.; Staehler, P.; Beier, M.; Tautz, D., Tests of rRNA hybridization to microarrays suggest that hybridization characteristics of oligonucleotide probes for species discrimination cannot be predicted. *Nucleic Acids Res.* **2006**, 34, e66.
6. Bhattacharya, P.K.; Cha, J.; Barton, J.K., 1H NMR determination of base-pair lifetimes in oligonucleotides containing single base mismatches. *Nucleic Acids Res.* **2002**, 30, 4740-50.
7. Lockhart, D. J.; Winzeler, E.A., Genomics, gene expression and DNA arrays. *Nature* **2000**, 405, 827-36.