University of Southampton

A Mass Spectrometry Based Hybridisation Assay for Single Nucleotide Polymorphism Analysis

Rachel Jennifer Ball

Doctor of Philosophy

Faculty of Engineering, Science and Mathematics

School of Chemistry

January 2005

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF ENGINEERING, SCIENCE AND MATHEMATICS
SCHOOL OF CHEMISTRY

Doctor of Philosophy

A MASS SPECTROMETRY BASED HYBRIDISATION ASSAY FOR SINGLE NUCLEOTIDE POLYMORPHISM ANALYSIS

By Rachel Jennifer Ball

Peptide nucleic acid is an analogue of deoxyribonucleic acid in which the phosphate backbone has been replaced with an achiral, uncharged polyamide backbone. This modification allows peptide nucleic acids to form duplexes with high affinity and sequence specificity with complementary deoxyribonucleic acid strands.

An approach that uses peptide nucleic acid probes in a hybridisation assay to detect single nucleotide polymorphisms has been developed. The probes hybridise to an immobilised single stranded polymerase chain reaction product and the peptide nucleic acid/deoxyribonucleic acid/bead conjugate is analysed directly by matrix-assisted laser desorption/ionisation-time of flight mass spectrometry. The genotype of the deoxyribonucleic acid is determined by the molecular mass of the peptide nucleic acid probe released. From one peptide nucleic acid scaffold four different types of peptide nucleic acid probes have been synthesised: an unmodified probe, a charge tagged probe, a probe with charge tagged mass marker and a probe with isotopically labelled charge tagged photo-cleavable mass marker.

Cystic fibrosis mutation analysis was chosen to evaluate the peptide nucleic acid hybridisation assay; in particular, studies have focused on mutations on the W1282X locus. The peptide nucleic acid probes have shown discrimination against mismatches and the neutral backbone does not form cation adducts like conventional deoxyribonucleic acid probes, which can hamper the sensitivity and resolution of detection by mass spectrometry. All four types of peptide nucleic acid probe have been successfully used in the hybridisation assay to give unambiguous detection of point mutations.

To ‘My Mum’

Contents

List of Figures

List of Tables

List of Schemes

Declaration

I, Rachel Jennifer Ball, declare that the thesis entitled "A Mass Spectrometry Based Hybridisation Assay for Single Nucleotide Polymorphism Analysis" and the work presented in it are my own. I confirm that:

- this work was done wholly or mainly while in candidature for a research degree at this University;
- no part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution;
- where I have consulted the published work of others, this is always clearly attributed;
- where I have quoted from the work of others, the source is always clearly given. With the exceptions of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;

Rachel Jennifer Ball, January 2005

Acknowledgements

I would like to express my gratitude to my supervisors, Dr John Langley and Prof. Tom Brown, for their ideas and assistance throughout the last three years, without which my PhD would not have been possible.

I would like to thank Syngenta, not only for the generous financial support of this research but for the continued advice and assistance of Mr Peter Massey, Dr Philip Green and Dr Peter Kilby.

Special thanks go to Dr Rohan Ranasinghe, Dr Lee Patient, Dr Sadie Osborne, Miss Rachel Downes and Miss Naomi Hammond for their time-spent proof reading this document. I would also like to thank Dr Dorcas Brown for assistance preparing the PNA probes and Ms. Julie Herniman for her advice and technical assistance. I thank Dr Mark Dixon and Dr Lee Patient for their advice for helpful discussions. I also thank all the past and present members of the Brown research group and mass spectrometry group I have had the pleasure of working along side over the past three years.

In addition, I thank Mrs Joan Street and Dr Neil Wells for all their assistance and providing excellent NMR resources.

Finally, I would like to thank Lee and my other close friends (Sarah, Rachel, Dawn and the past, present and associate members of the Shirley house) for their support throughout my PhD. I would particularly like to thank my mum and sister for their love and continued encouragement during the last year.

Abbreviations

δ	chemical shift (parts per million)
ν_{max}	Infrared absorption maximum
μ	micro
+	positive
-	negative
^{1}H	proton (NMR)
^{13}C	carbon (NMR)
°C	degrees celsius
A	adenine
ABI	Applied Biosystems
Ala	alanine
BOC	*tert*-butoxycarbonyl
Bhoc	benzhydryloxycarbonyl
br	broad
BSA	bovine serum albumin
C	cytosine
c	centi
CA	collisionally activated
CBZ	benzyloxycarbonyl
CD	circular dichroism
CF	cystic fibrosis
CFTR	cystic fibrosis transmembrane conductance regulator
CI	chemical ionisation
CID	collision induced dissociation
C_5Q	6-[(2,5-dioxotetrahydro-1*H*-1-pyrrolyl)oxy]-6-oxohexyl(trimethyl)ammonium bromide
dATP	2'-deoxyadenosine 5'-triphosphate
DEAD	diethylazodicarboxylate

DIAD	diisopropylazodicarboxylate
ddNTP	2',3'-dideoxynucleoside 5'-triphosphate
dGTP	2'-deoxyguanosine 5'-triphosphate
DIPEA	*N,N*-diisopropylethylamine
DME	1,2-dimethoxyethane
DMF	*N,N*-dimethylformamide
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
DNsCl	2,4-dinitrosulfonyl chloride
dNTP	2'-deoxynucleoside 5'-triphosphate
EDC	*N*-ethyl-*N*'-(3-dimethylaminopropyl)carbodiimide hydrochloride
EI	electron ionisation
eq.	equivalents
ES	electrospray ionisation
EtBr	ethidium bromide
EtOAc	ethyl acetate
f	femto
FAB	fast atom bombardment
Fmoc	9*H*-9-fluorenylmethoxy)carbonyl
FRET	fluorescence resonance energy transfer
FT-ICR	Fourier transform-ion cyclotron resonance
G	guanine
Gly	glycine
h	hour
HOBt	*N*-hydroxybenzotriazole
3-HPA	3-hydroxypicolinic acid
HPLC	high-performance liquid chromatography
HRMS	high-resolution mass spectrometry
Hz	hertz
k	kilo
L	litre
LRMS	low resolution mass spectrometry

Lys	lysine
M	Molar
m	medium
MALDI	matrix-assisted laser desorption/ionisation
MHz	mega hertz
min	minutes
mm	millimetre
mol	moles
MS	Mass spectrometer
m/z	mass to charge ratio
n	nano
NAP	nucleic acid purification
ND	not determined
NMM	4-methylmorpholine
NMO	4-methylmorpholine *N*-oxide
NMR	nuclear magnetic resonance
NsCl	2-nitrosulfonyl chloride
O.D.	optical density at 260 nm
p	phosphate
PA	picolinic acid
PCR	polymerase chain reaction
Ph	phenyl
Phe	phenylalanine
pmol	picomole
PNA	peptide nucleic acid
POA	peptide organic acid
PSD	post source decay
RF	radio frequency
R_f	retention factor
RP-HPLC	reverse phase high-pressure liquid chromatography
rt	room temperature (25 °C)
RT	real time

s	strong
SAP	shrimp alkaline phosphatase
sat.	saturated
SNP	single nucleotide polymorphism
SORI-CID	sustained off-resonance irradiation-collision induced dissociation
Su	succinimide
T	thymine/thymidine
T	tesla
Taq	*Thermus aquaticus*
TBE	tris-borate-ethylenediamine tetraacetic acid
TEAA	triethylammonium acetate
tert	tertiary
TFA	trifluoroacetic acid
THF	tetrahydrofuran
TLC	thin layer chromatography
T_m	melting temperature
TOF	time of flight
TPAP	tetrapropylammonium perruthenate
Tris	tris(hydroxymethyl)aminomethane
Trt	trityl
SDS	sodium dodecyl sulfate
U	units of enzyme
UV	ultraviolet
V	volts
w	weak

1. Introduction

1.1. Nucleic Acids

1.1.1. Primary Structure

Deoxyribonucleic acid (DNA) is a polymeric molecule composed of monomer units of deoxyribonucleotides. The constituents of each deoxyribonucleotide, shown in **Figure 1.1** are a pentose sugar, a heterocyclic base and a phosphate moiety. The 5'-hydroxyl group of one deoxyribose unit is covalently linked to the 3'-hydroxyl group of the subsequent unit *via* a phosphate group (phosphodiester linkage). The four bases, either functionalised purines (adenine (**A**) and guanine (**G**)) or functionalised pyrimidines (thymine (**T**) and cytosine (**C**)) are attached to the C1 atom of the pentose sugar by means of a glycosidic linkage (**Figure 1.2**).

The variation in DNA arises from the order of the four bases (**A**, **G**, **C** and **T**). Convention denotes that the DNA sequence is recorded from the 5'- to the 3'-direction *i.e.* from the 5'-terminus to the 3'-terminus using the single letter abbreviations of the different nucleobases. It is important to note that at physiological pH each phosphate is deprotonated giving DNA an overall negatively charged backbone.

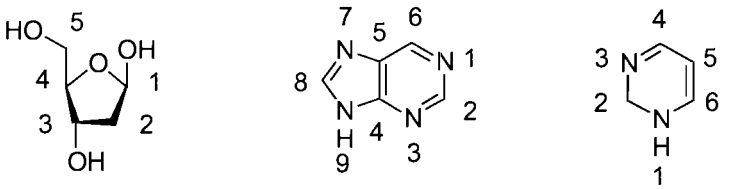


Figure 1.1. The nomenclature of deoxyribose and the nitrogenous bases. Purine, a bicyclic structure and pyrimidine, a monocyclic base.

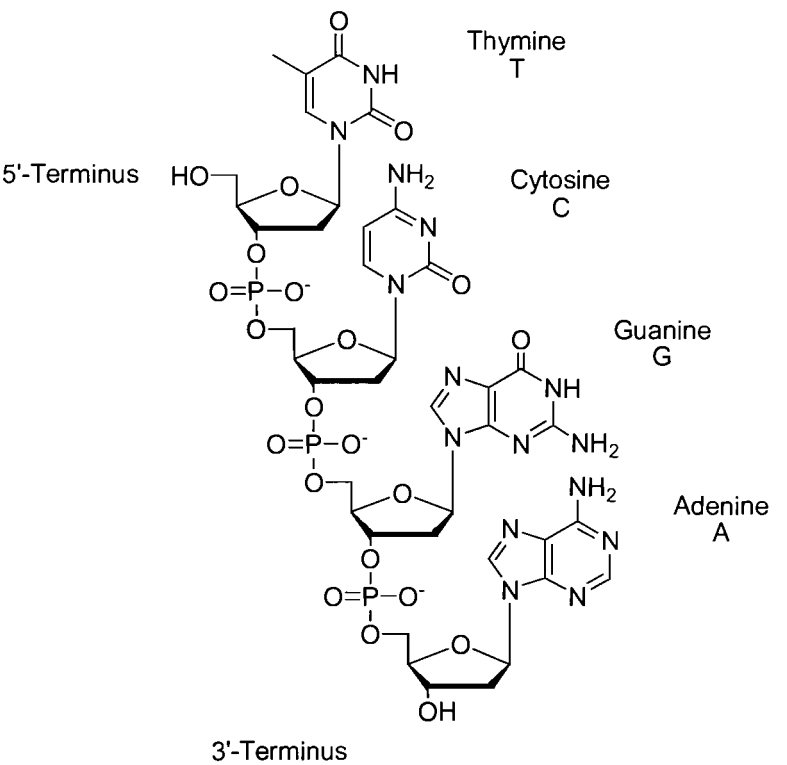


Figure 1.2. The structure of the sugar phosphate backbone and the four bases found in DNA.

1.1.2. Secondary Structure

The three dimensional structure of DNA was determined in 1953 by James Watson and Francis Crick from X-ray diffraction photographs of DNA fibres obtained by Rosalind Franklin and Maurice Wilkins.[1] Watson and Crick derived a structural model, which described two polynucleotide chains organised in a duplex.[2, 3] The coiling results in a right handed helix where the two complementary strands run in opposite directions. The purine and pyrimidine bases are orientated on the inside of the helix, perpendicular to the axis of the helix, while the phosphate and deoxyribose units lie on the outside forming the external backbone as described in **Figure 1.3**.

The two polynucleotide chains are held in place by hydrogen bonds between a purine base and a pyrimidine base where **A** pairs with **T** with two hydrogen bonds, and **C** pairs with **G** with three hydrogen bonds (**Figure 1.4**). This structure further confirmed the findings of Chargaff who found that the ratio of **A** to **T** and **G** to **C** in DNA was nearly always 1:1. However, the ratio of **G**:**C** to **A**:**T** varied from species to species. [4, 5] In this way, the sequence of one strand directly defines the sequence of the second strand immediately suggesting a mechanism for the replication of DNA.

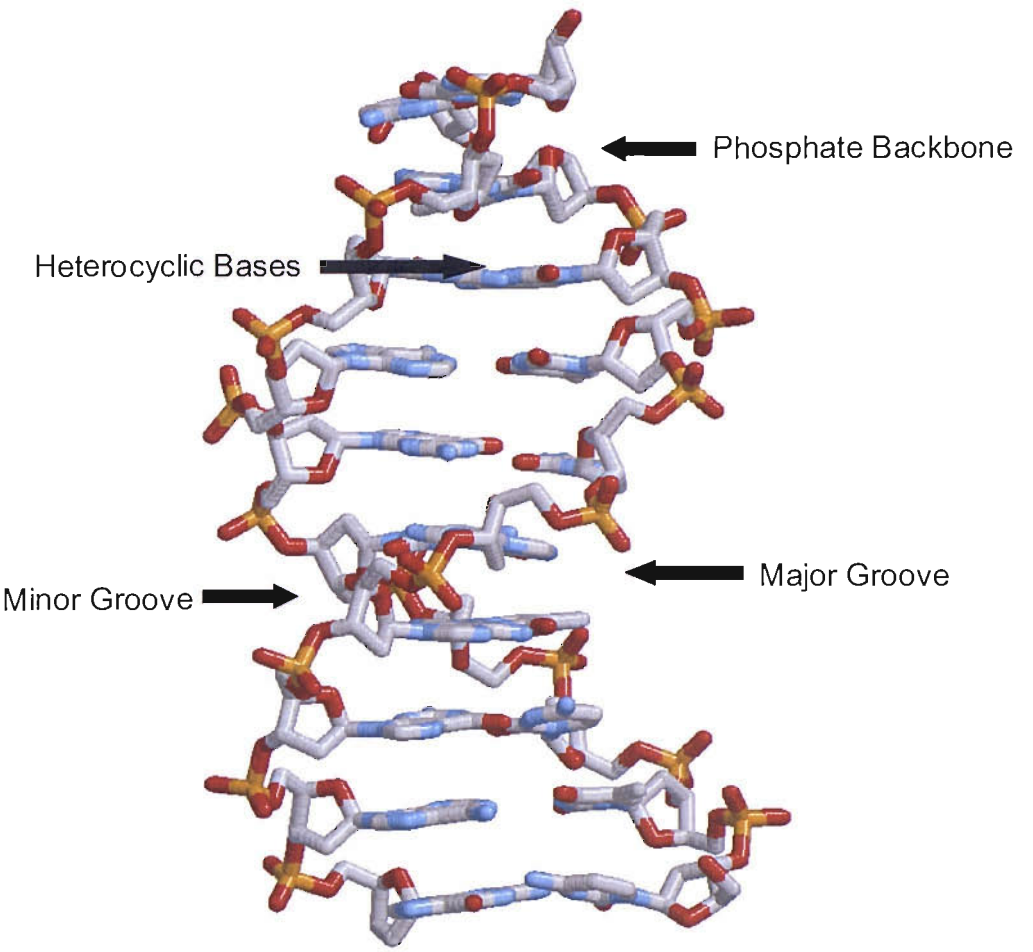
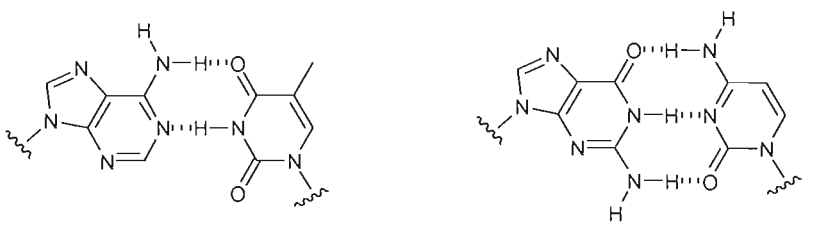


Figure 1.3. A crystal structure of 5'-d(CTCTCGAGAG) obtained by X-ray diffraction showing the minor, major grooves and the secondary structure of DNA.



A : T G : C

Figure 1.4. The complementary base pairs or Watson Crick base pairs found in DNA. The adenine-thymine base pair has two hydrogen bonds and the guanine-cytosine base pair has three hydrogen bonds.

1.2. Single Nucleotide Polymorphisms

The human genome contains the set of genetic instructions that describes an individual. It consists of 23 pairs of chromosomes containing an estimated 3 $\times 10^4$ genes that are described by the sequence of 3 billion DNA bases.

Single nucleotide polymorphisms or SNPs are DNA sequence variations that occur when a single nucleotide **A**, **G**, **C** or **T** is variable *i.e.* when sequence alternatives, known as alleles exist. In humans, SNPs are usually biallelic polymorphisms, that is, the nucleotide identity at the polymorphic site is generally constrained to two alternatives rather than the four possibilities that in principle can occur.[6] For a variation to be considered a SNP it must occur at a frequency of greater than 1 % of the population. SNPs make up 90 % of all human genetic variation occurring every 1.91 kilobases in the human genome. The remaining variations are attributed to insertions or deletions of one or more bases, repeat length polymorphisms and rearrangements.[7] At the end of the 1990s a large effort was made by the SNP Consortium Ltd. to identify up to 300,000 SNP markers distributed evenly throughout the human genome providing public genomic data.

Whilst only a small fraction of SNPs lie within the coding region of genes and an even smaller proportion are responsible for alterations to the amino acid expressed in proteins, they may be in close proximity to deleterious mutations on the chromosome. The wide distribution and frequency of these variations make SNPs an extremely valuable tool for determining identity, establishing genetic linkages or indicating an individual's predisposition to a disease.[8] In addition to their high frequency SNPs are stable over generations so are not susceptible to spontaneous change.

1.3. Polymerase Chain Reaction

In order to effectively detect SNPs it is necessary to amplify a region of DNA to produce a suitably intense signal to either confirm the presence or absence of the SNP. In 1979 Khorana[9] and co-workers described an approach for replicating a region of DNA duplex using two primers. However, the concept of using this approach repeatedly in amplification was not realised until Kary Mullis introduced the polymerase chain reaction (PCR) in 1984.[10-12] In a simple buffer system a short region of template DNA is copied by a DNA polymerase using deoxynucleotidetriphosphates (dNTPs) as the building blocks for the new strand. Sequence-specific oligonucleotide primers, designed to bind to the template, define the region of the DNA to be amplified. The strands of the DNA template are separated or denatured by heat, causing the hydrogen bonds to sever (**Figure 1.5**). The complementary synthetic oligonucleotide primers anneal to the single stranded DNA template and the polymerase begins to add dNTPs to the 3'-hydroxyl group of the primers reforming a region of duplex, a process known as extension. On heating, the newly formed duplex denatures and each single strand (both the original template and the newly synthesised strand) serves as a template for further DNA replication. In this way, an exponential increase in the number of copies of the target DNA sequence is observed as the three steps, denaturation, annealing and extension are repeated.

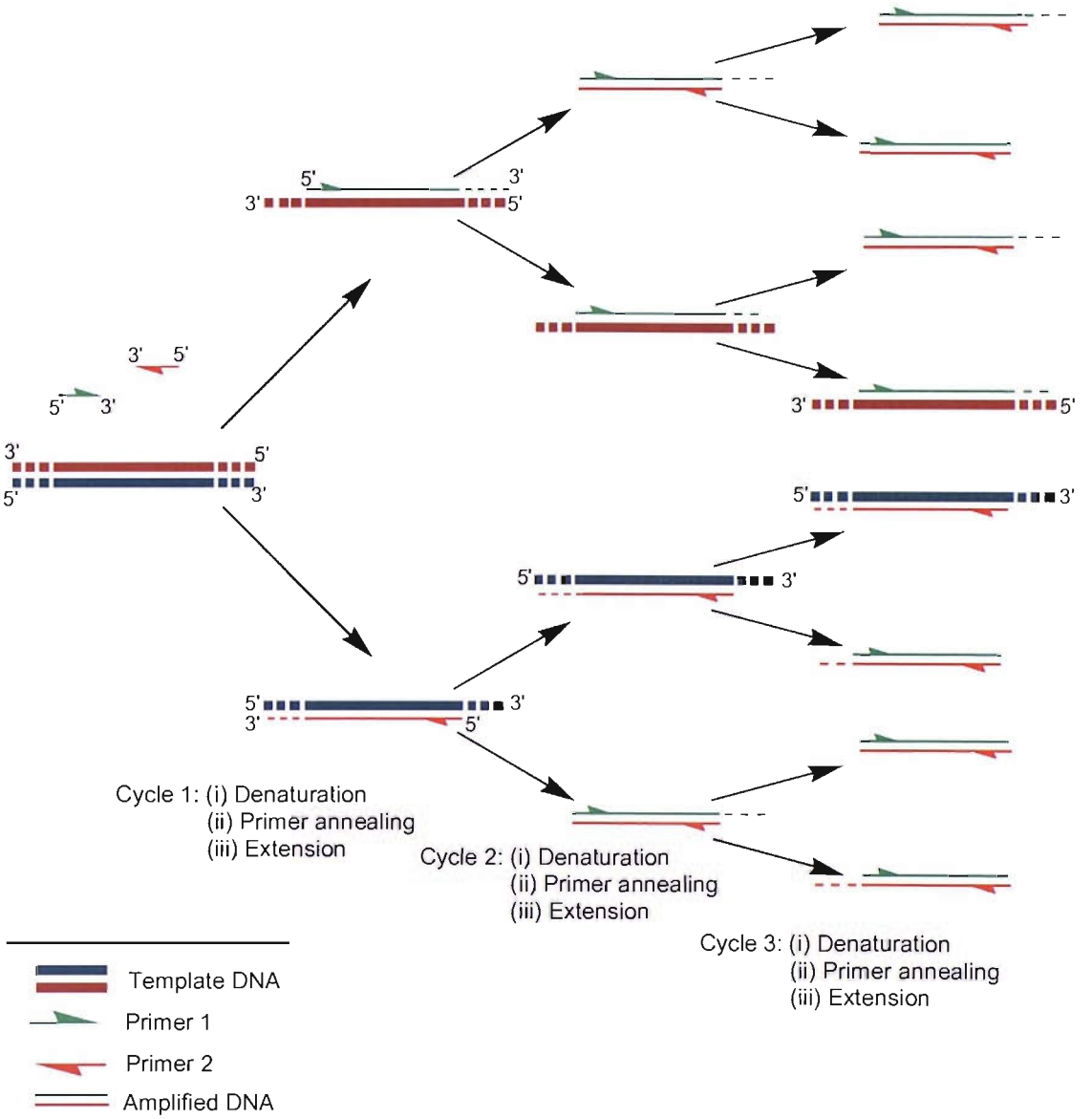


Figure 1.5. A schematic diagram of the first three cycles of a PCR reaction.

1.4. Ultraviolet-Melting

Ultraviolet (UV)-melting is used to determine the thermodynamics of binding for nucleic acids and is frequently used to establish the melting temperature or T_m of an oligonucleotide and its complement. This information is essential when designing experiments for applications such as PCR and hybridisation assays. The common method to experimentally determine the T_m is to use a thermostatted cell in a UV spectrometer.

The nucleobases are not exposed to the aqueous environment when in a duplex due to the electronic interaction with neighbouring bases. These interactions are known as base stacking. As a consequence the extinction coefficient of the double stranded DNA (at 260 nm) is typically 40 % less than the extinction coefficient of the two single strands. For this reason the absorbance of a buffered solution containing two oligonucleotides is measured at 260 nm. At low temperatures the two oligonucleotides exist as a duplex. When the solution is heated the hydrogen bonds between the bases break and base stacking is disrupted leading to an increase in absorbance. A plot of the absorbance at 260 nm as a function of temperature gives rise to a melting curve (**Figure 1.6**). The T_m is defined as the temperature at which 50 % of the oligonucleotide and its complement are in the duplex form and can be calculated for the maximum point of the first derivative of the melting curve (**Figure 1.6**).

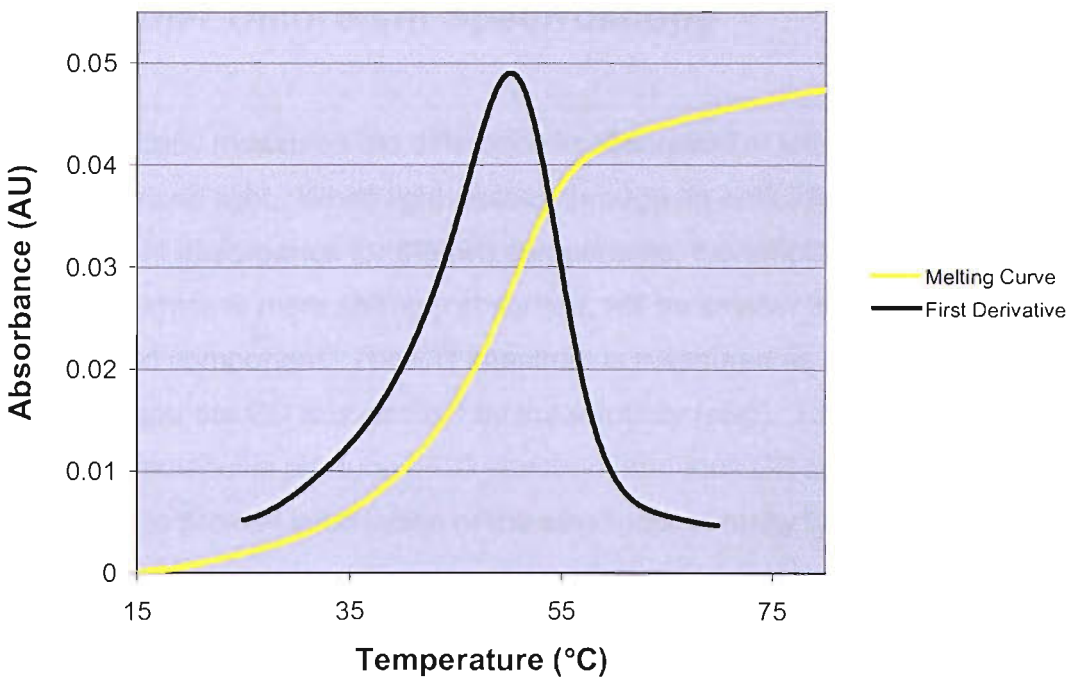


Figure 1.6. A melting curve of a DNA/PNA duplex from 15 °C to 80 °C. The smoothed first derivative of the melting curve is shown.

1.5. Circular Dichroism Spectroscopy

CD spectroscopy measures the difference in absorption of left and right handed polarised light. When light passes through an optically active sample with a different absorbance for the two components, the amplitude of the component, which is more strongly absorbed, will be smaller than that of the less absorbed component. The CD spectrum is measured as a function of wavelength and the CD is quantified by the ellipticity (ellip). Chiral or asymmetric molecules produce a CD spectrum and thus CD spectroscopy can be used to provide information of the structures of many types of biological macromolecules. Proteins and nucleic acids contain optically active elements and adopt different types of three-dimensional structures and thus exhibit distinct CD signals. The absence of regular structure results in zero CD intensity, while ordered structure results in a spectrum, which can contain both positive and negative signals. The wavelengths of light that are most useful for examining the structures of proteins and DNA are in the UV and vacuum UV (VUV) (160 – 300 nm). This is the region of the electronic transitions of the peptide backbone and the purine and pyrimidine bases in DNA. In biology, CD spectroscopy is typically used to study secondary structure in proteins and nucleic acids and access the thermal stability by following changes in the CD spectrum with increasing temperature. It has also been used to study protein-nucleic acid and protein-protein interactions and to determine the stoichiometry of such interactions. CD is based on measurements in the solution phase and thus allows simple and quick experiments without extensive sample preparation.

1.6. Mass Spectrometry

The history of mass spectrometry started with the work of J. J. Thompson in the early twentieth century who put forward the principle ideas underlying mass spectrometry.[13] A mass spectrometer is an instrument, which primarily measures the mass to charge ratio or m/z of ions. The different functional units of a mass spectrometer are described in **Figure 1.7**. Different technologies have been introduced for each of these units and the method of choice depends on the chemical properties of the sample being analysed. The sample is introduced through the inlet and is ionised in an ionisation source. Once formed, the ions are extracted from the ion source and separated according to their mass to charge ratio (m/z) in one or several mass analysers. The ion abundance is measured as the ions reach the detector and the information is relayed to the data system where it is manipulated into a plot of the m/z ratio with respect to ion abundance *i.e.* a mass spectrum. The hybridisation assay described in this project was primarily designed for detection using a matrix-assisted laser desorption/ionisation-time of flight mass spectrometer (MALDI-TOF-MS) and all samples in this project were analysed by this technique. For this reason only the principles of the MALDI-TOF-MS will be discussed in this introduction.

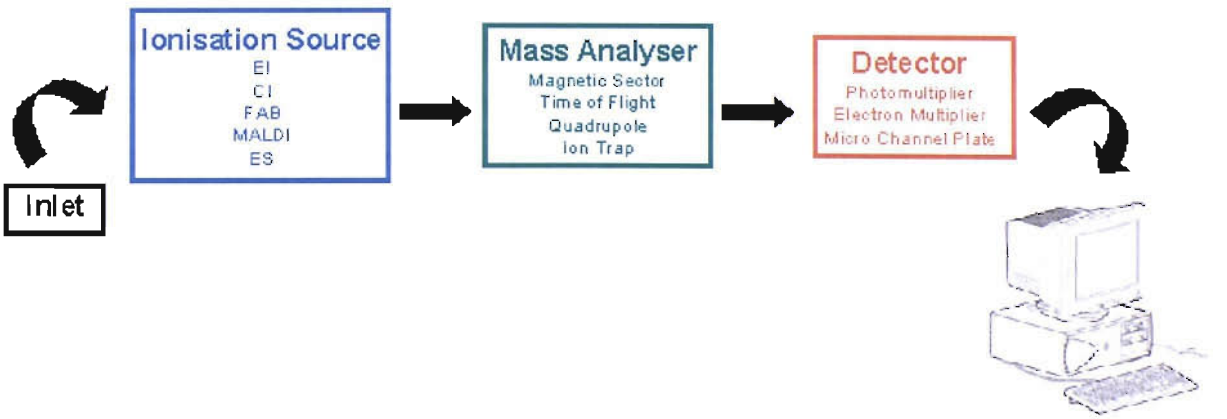


Figure 1.7. Functional components of a mass spectrometer.

1.6.1. Matrix-Assisted Laser Desorption/Ionisation

Lasers originally found applications in mass spectrometry in the early 1970s for desorption and ionisation of small organic compounds that contained chromophores.[14] The abrupt energy absorption initiates a change in phase of the absorbing sample whilst also inducing ionisation. However, it was not until the late 1980s when the breakthrough came with the realisation that incorporation of an analyte into the crystalline structure of a small UV absorbing compound (the matrix) provided a vehicle for intact ions to be created from larger molecules.[15-17] The function of the matrix is to absorb the laser energy thus desorbing the matrix and sample from the surface of the target. This results in the transfer of the analyte into the gas phase without thermal stress.

As described in **Figure 1.8**, a solution of the sample and matrix is deposited on the MALDI target and allowed to co-crystallise as the solvent evaporates. A pulsed laser beam irradiates the matrix, which absorbs the energy from the light. This energy increases the internal energy of the matrix causing the matrix to melt and vaporise. During this process the analyte molecules, which are embedded in the matrix, are desorbed. Despite vigorous efforts by researchers to understand the underlying processes of ionisation the absolute MALDI process is still unclear[18-21] and it is probable that there is no single mechanism to explain all of the ions generated. One approach suggests the gas phase protonation or deprotonation following desorption. After desorption, in the gas phase a chemical process occurs in which a proton is transferred to the sample from the matrix or is removed from the sample resulting in the formation of positive and negative sample ions respectively (**Figure 1.8**).

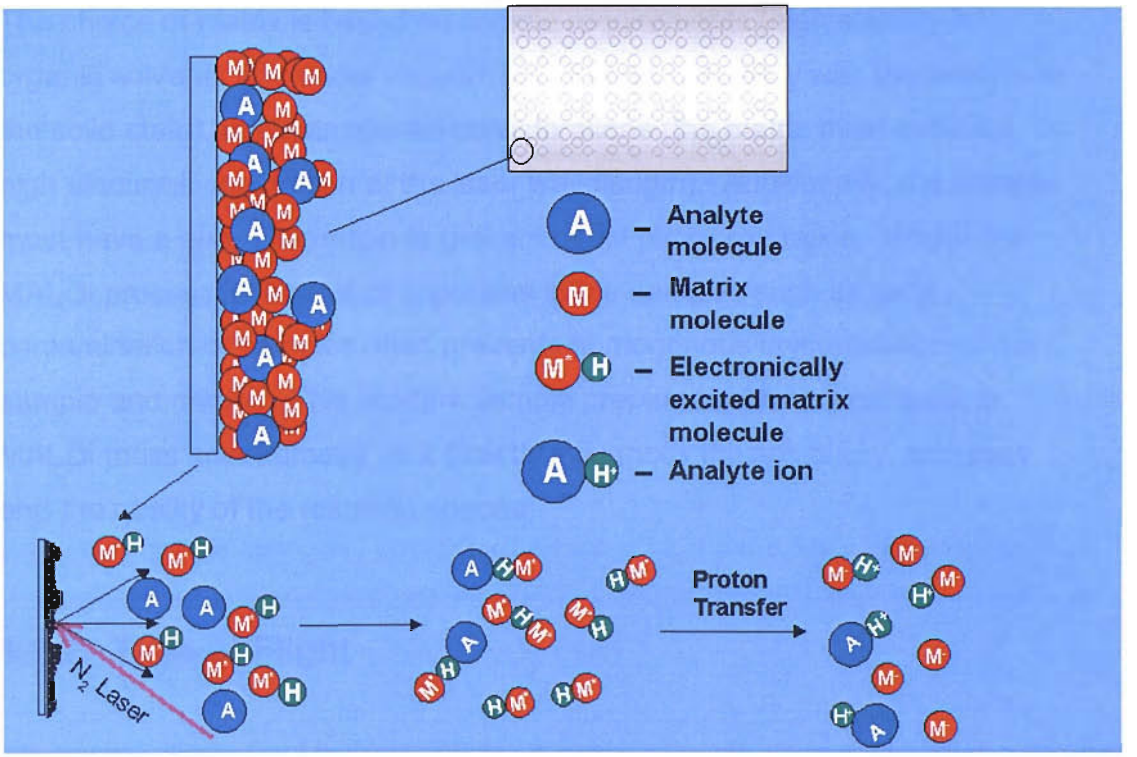


Figure 1.8. MALDI ionisation process. The matrix and analyte are deposited on the target. After irradiation by the laser the sample is desorbed from the target and ionised. Reproduced with permission from Kathryn Lilley, University of Cambridge.

This mechanism allows the formation of both protonated matrix and analyte molecules, which are commonly observed in the mass spectrum. An alternative mechanism suggests the ions observed in the mass spectrum are already present in the solid phase and are only desorbed during the laser pulse. This condensed phase theory for ionisation has been supported by the work of Zhu and co-workers who observed different detection efficiencies for peptides with basic amino acid residues.[22]

Lasers of wavelengths in the near UV are used in UV-MALDI, most commonly, 337 nm nitrogen and 355 nm Nd-YAG lasers however lasers emitting in the infrared (IR) region are used in IR-MALDI.

The choice of matrix is based on sample requirements (high stability in organic solvents and under vacuum and a good miscibility with the analyte in the solid state) and instrumental considerations (the matrix must exhibit a high electronic absorption at the laser wavelength). Additionally, the sample must have a poor absorption to prevent direct photo-excitation. Whilst the MALDI process is tolerant of impurities in the samples such as salts, contamination of samples often prevents homogenous crystallisation of the sample and matrix. This renders sample preparation the critical point in MALDI mass spectrometry as it directly influences the sensitivity, accuracy and the quality of the resulting spectra.

1.6.2. Time of Flight

The MALDI ionisation method is ideally coupled to the TOF mass analyser because of the pulsed nature of the desorption laser. As a result all ions produced from a single laser shot can be simultaneously separated by the TOF mass analyser. Additionally, in principle, the TOF analyser has no upper mass limit so is well suited to softer ionisation methods that promote intact large molecules into the gas phase. The TOF mass analyser has a greater efficiency and higher sensitivity than scanning mass spectrometers such as sectors or quadrupoles. These types of instruments require a continuous stream of sample ions to scan the desired mass range. At any given point, only ions of a certain mass have a stable trajectory and proceed through the mass analyser reaching the detector. As a result all other ions are lost. The additional virtues of the TOF are the simplicity of the instrument design and the low cost of manufacture and maintenance.

1.6.2.1. Linear Time of Flight Mass Analyser

Whilst the linear TOF mass analyser was first introduced in 1946 by Stephens[23] the first publication describing a TOF mass analyser appeared later in 1953.[24] The TOF mass analyser operates on the principle that when a defined population of ions, of differing m/z, are subjected to an electric field and are subsequently allowed to drift in a field free region the ions will transverse the region in a time which is dependant on the m/z ratio. Low mass ions have a greater velocity and a shorter flight time than heavier mass ions when singly charged. Additionally, highly charged species have shorter drift times than singly charged species. A schematic illustration of a simple TOF mass analyser is shown in **Figure 1.9**.

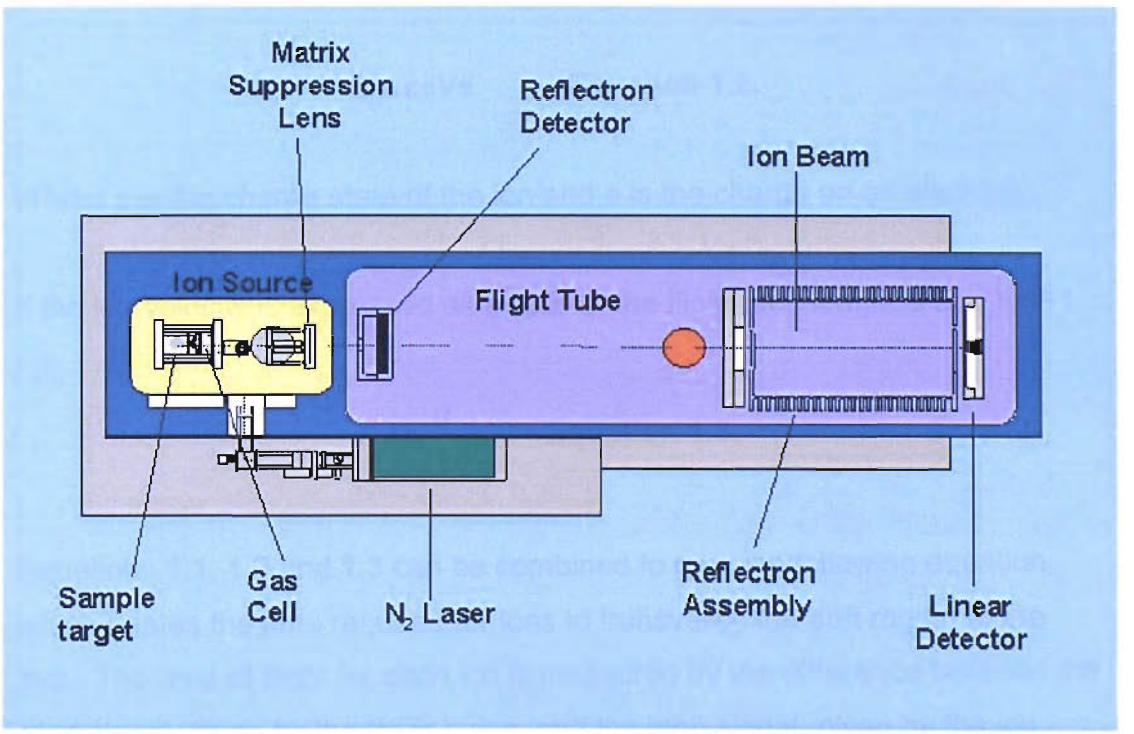


Figure 1.9. Schematic illustration of a TOF mass analyser. The main components of the TOF mass analyser are labeled. Reproduced with permission from Kathryn Lilley, University of Cambridge.

At intervals ions are subjected to a high voltage pulse typically in the region of 20-30 kV. The voltage is either positive or negative depending upon the selection of negative or positive ion detection. This accelerates ions into the field free flight tube. The ions traverse the region of distance **d** between the source and the detector. The ions with mass **m** and charge **q** have kinetic energy **E** that can by described be the following equation

$$E = \frac{1}{2}mv^2 \qquad \textbf{Equation 1.1.}$$

Where **v** is the velocity of the ion.

The kinetic energy **E** can also be expressed with respect to the potential applied (**Vs**) to the ion according the following equation

$$E = zeVs \qquad \textbf{Equation 1.2.}$$

Where **z** is the charge state of the ion and **e** is the charge on an electron.

If the ion velocity is expressed as a ratio of the flight tube length **d** and time **t**

$$v = \frac{d}{t} \qquad \textbf{Equation 1.3.}$$

Equations **1.1, 1.2** and **1.3** can be combined to give the following equation, which relates the time required for ions to transverse the drift region to the m/z. The time of flight for each ion is measured by the difference between the start signal, given by the laser pulse, and the stop signal, given by the ion impinging on the detector.

Since the potential **Vs** and length of drift tube **d** are known and **e** is constant the m/z can be calculated directly from the time of flight using **Equation 1.4**.

$$t = d\sqrt{\frac{m}{2zeVs}} \qquad \textbf{Equation 1.4.}$$

After a short time delay to ensure all ions have reached the detector the next pulse of ions are introduced into the flight tube and a mass spectrum is generated from the sum of several laser pulses.

One of the drawbacks of the linear TOF mass analyser is the poor mass resolution. The resolution of a mass spectrometer is the ability to distinguish between ions of different m/z values. It can also be used to measure the performance of a mass analyser. The resolution describes the degree to which two ions of adjacent m/z can be distinguished. There are many definitions, depending on specific applications and instrument types being used. One common definition for resolution **R**, is the ratio of the mass of an ion (M_1) to the difference in mass between M_1 and the adjacent ion of higher mass (M_2) (**Equation 1.5**).

$$R = \frac{M_1}{M_2 - M_1} \qquad \textbf{Equation 1.5.}$$

One additional definition of resolution based on the width of the ion, is the full width of the ion at half its maximum height (FWHM). For a single ion of mass **M** the resolution can be expressed as a function of the width of the ion at 50 % of the maximum ion height (ΔM) (**Equation 1.6**)

$$R = \frac{M}{\Delta M} \qquad \textbf{Equation 1.6.}$$

The mass resolution is affected by factors that create a distribution in flight times of ions with the same m/z. This includes variation of the initial kinetic energies[25] and a distribution of the initial spatial spread of same m/z ions. To date two solutions have been used to overcome this problem; the reflectron time of flight mass analyser and delayed pulsed extraction.

1.6.2.4. Reflectron Time Of Flight Mass Analyser

A primary contribution to the loss in mass resolution is attributed to the range in kinetic energies of individual m/z ions as they leave the ion source. In order to overcome this problem the reflectron was introduced. First proposed by Mamyrin[26, 27] the reflector, located at the end of the flight tube is composed of a series of grids and ring electrodes. Ions with a higher kinetic energy are first to arrive at the reflectron. These ions penetrate deeper thus have a longer flight path in the reflectron than ions with less kinetic energy and therefore arrive at the detector at the same time. As the ions emerge from the reflectron the ion direction is reversed to send the ions towards the detector adjacent to the source (as shown in **Figure 1.9**). Consequently, any differences in energy of ions of the same m/z value are dissipated resulting in these ions being focused hence arriving at the detector at the same time. This correction leads to an increase in mass resolution for all stable ions in the spectrum.

In addition to enhancing mass resolution, the reflectron permits tandem mass spectrometry experiments, such as post source decay (PSD), which can be exploited to obtain structural information. PSD involves the detection of fragments generated from the intact ion (product ion) in the field free region between the source and the detector.

Ions that fragment in the field-free region of the drift tube retain essentially the same velocity as intact ions and cannot be distinguished from stable ions in a linear TOF instrument. Fragments are generated either by an excess of energy at ionisation or by collisions with neutral species (collision-induced dissociation (CID) or collision-activated dissociation (CAD)). CID is the process in which an ion is fragmented as a result of a collision with a neutral species. Collisions can be random such as with residual gas molecules or neutral matrix ions or controlled using a gas collision cell. During the collision, part of the translational energy of the ion is transferred to internal energy bringing the ion into an excited state, which leads to fragmentation.

1.6.2.3. Delayed Extraction

Whilst the reflectron compensates for the energy spread of the created ions, delayed extraction corrects any spatial differences that occur during the ionisation event. Improvements in mass resolution have been achieved by the use of delayed extraction (also referred to as pulsed ion extraction, pulsed extraction or dynamic extraction).[28, 29]

Delayed extraction is based on similar principles to time lag focusing[30] and was first used on a commercial instrument by Vestal and co-workers in 1995.[28, 31] In contrast to measurements where the ions are desorbed and directly accelerated into the drift tube, a short time delay is introduced after the ionisation event. During this period ions are allowed to drift in a region with no electric field resulting in the dispersion of the ion packet. The electric field is then applied and ions that have traveled further from the MALDI target receive less energy and as a consequence experience less acceleration than ions that remain in the source region for a longer period of time. This time delay (typically 1-10 μseconds) compensates for any initial spatial distribution amongst ions of the same m/z value, effectively focusing the ion packet and resulting in a much-improved performance.

One additional benefit of this technology is that any neutral species created during desorption are removed thus reducing ion-neutral collisions. In this way the initial desorption and ion formation process is separated from the acceleration step and therefore does not directly contribute to peak broadening and a loss in mass resolution.

1.7. Modern Techniques for Genetic Analysis

1.7.1. Fluorescence Based Methods for Detection of Genetic Mutations

Fluorescenct molecules absorb light of a specific wavelength (excitation wavelength) and after a period re-emit it at a longer wavelength (emission wavelength) of a lower energy.[32] Since the electronic structures of these molecules vary depending on the constituent groups and elements, the specific wavelength of excitation and emission can be tuned with synthetic manipulation of fluorescent dyes. Incorporation of these types of molecules in nucleic acids permits their detection.

1.7.1.2. Sanger Sequencing

Sanger sequencing is used to determine the sequence of nucleic acids of unknown composition. Also known as dideoxy sequencing, Sanger sequencing involves the generation of a number of labelled DNA fragments which differ in length but are a fraction of the entire sequence.[33] In each reaction a primer anneals to the template strand and is extended in the presence of 2'-deoxynucleoside 5'-triphosphates (dNTPs) and 2',3'-dideoxynucleoside 5'-triphosphates (ddNTPs).

Since the ddNTPs contain no 3'-hydroxyl group after incorporation of the ddNTP the primer cannot be extended any further. The fragments generated are then separated by gel electrophoresis. There are two popular methods of Sanger sequencing.

Dye-Terminator Sequencing

As described in **Figure 1.10** four ddNTPs each carrying a different fluorophore are added to the reaction mixture. This permits the identification of the terrninating base of each fragment.

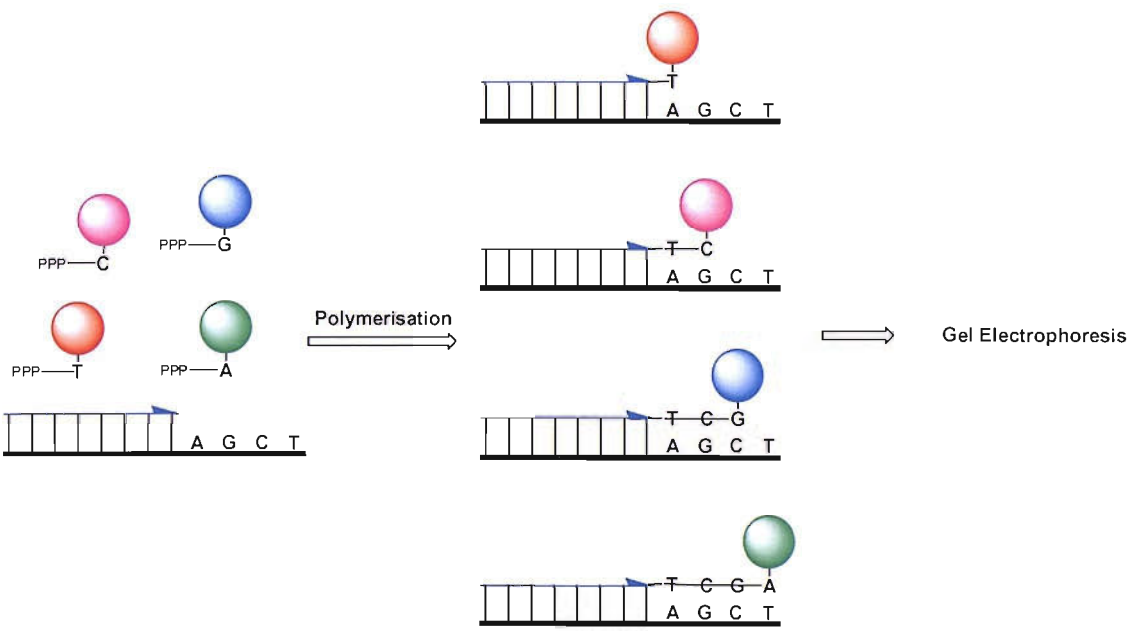


Figure 1.10. Dye-terminator Sanger dideoxy sequencing.

Dye-Primer Sequencing

In four independent reactions, each containing one ddNTP and three dNTPs, extension is carried out with a fluoresently labelled primer.[34]

1.7.1.2. Real Time-PCR

In real time (RT)-PCR amplification is monitored as the PCR reaction occurs. Commonly, fluorescent techniques are used to monitor the accumulation of the PCR product. RT-PCR can be used for genotyping and SNP detection.[35] Wild type, heterozygous and mutant genotypes give different amplification curves if suitable allele-specific primers are used. As well as the accumulation curve fluorescence melting curves can be obtained at the end of the PCR reaction. This in addition to speed and convenience makes RT-PCR an attractive technique.

1.7.1.3. TaqMan® Assay

This is a PCR-based assay in which the 5'- nuclease activity of the DNA polymerase, *Taq*, is utilised as well as polymerase properties.[36] The technique is based on the detection and accumulation of specific PCR products by hybridisation. The TaqMan® probe is an oligonucleotide, complementary to some region of the section of DNA flanked by the primers, which is labelled a fluorescent dye and quencher.[37] As the primer is extended by the polymerase, replication is obstructed by the TaqMan® probe. In this configuration the fluorophore is in a "dark" state *i.e.* is quenched *via* FRET. At the point where replication is obstructed the polymerase exploits its exonuclease activity and degrades the probe from the 5'- end. This results in separation of the fluorophore and quencher and FRET is essentially eliminated (**Figure 1.11**). This leads to an accumulation of the fluorescent signal observed from the released fluorophore over the cycles of the PCR, until the digestion reaches a plateau.[38] The TaqMan® technology has been used extensively as a tool in genetic analysis, detecting the Hepatitis C virus[39] and cancer susceptibility in humans.[40]

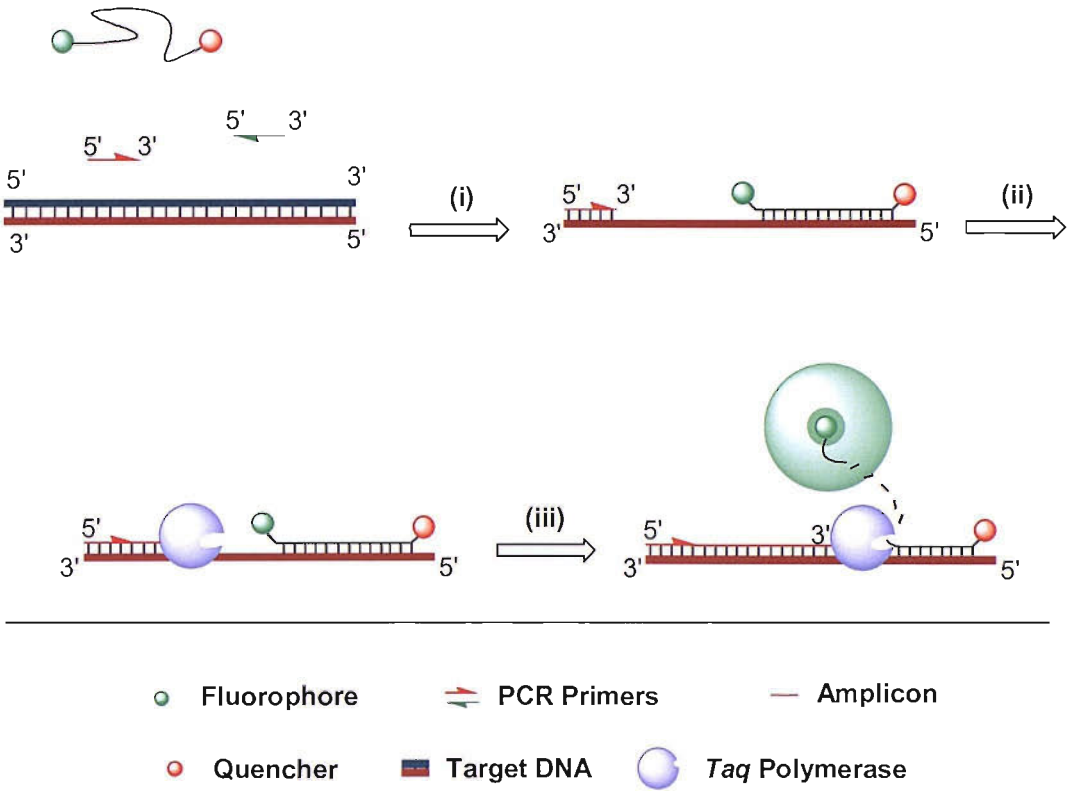


Figure 1.11. A schematic representation of the TaqMan® assay. i) Heat to denature genomic DNA. Cool to allow annealing of primers and TaqMan probe to target DNA. The TaqMan® probes fluorophore is quenched by FRET. ii) Extension of primer by polymerase in the 5'- to 3'- direction. iii) Exonuclease activity by *Taq* releases fluorophore.

1.7.1.4. Molecular Beacons

The molecular beacon technology is a hybridisation-based assay that utilises oligonucleotide probes that become fluorescent as a consequence of a conformational change. Molecular beacons posses a stem loop structure.[41] The beacon is designed so that at either end there is a small region of DNA, which is self-complementary, and consequently forms a double stranded stem. The fluorophore and quencher are covalently attached to either end of the stem. The probe region of the beacon complementary to the target DNA sequence forms the loop.

When the stem is formed, two dyes are held close together and fluorescent quenching occurs. When a complementary target is present the stem structure is broken and the fluorophore and quencher separated (**Figure 1.12**). This results in a fluorescent signal in the presence of a fully complementary target. By balancing the melting temperature of the stem with the most stable mismatch, the probe can be designed so that it only binds to a fully complementary target. Molecular beacons have found application in DNA in genotyping of human alleles.[42]

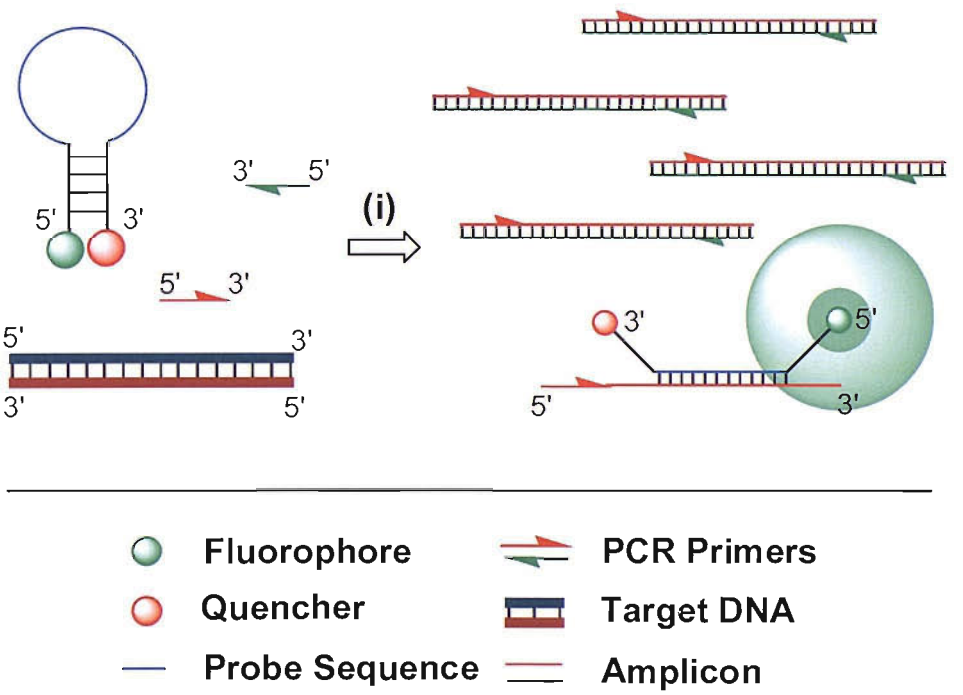


Figure 1.12. The structure and mechanism of action of a molecular beacon. i) PCR amplification of target DNA and annealing of molecular beacon to amplicon.

As with TaqMan®, the molecular beacon format is combined with PCR to produce a closed-tube assay in which the fluorescence can be monitored in real time as the PCR product accumulates. Different formats have been developed in an attempt to increase signal strength and accuracy, *e.g.* sunrise® primers,[43] scorpion primers[44] and duplex scorpion primers.[45]

1.7.1.5. DNA Micro Arrays

DNA micro arrays are composed of oligonucleotide probes complementary to the region of DNA of interest non-covalently bound to a solid surface. Hybridisation of a fluoresently labelled PCR product to the immobilised probe indicates the presence of a particular sequence in the amplification mixture. Positive results are detected by a fluorescent scanner. As described in **Figure 1.13** each spot indicates an area where an oligonucleotide has been attached. The colour of the spot indicates whether a complementary oligonucleotide carrying a fluorescent label has hybridised. White or yellow/green spots corresponding to a negative test and all the others are a positive result. Micro arrays are typically used for the simultaneous analysis of many genetic factors. However this methodology relies on the differential hybridisation stability of the target molecules that are perfectly base-paired to the immobilized probe *versus* the ones that contain one or more mismatches. For this reason the scope for multiplexing is limited.

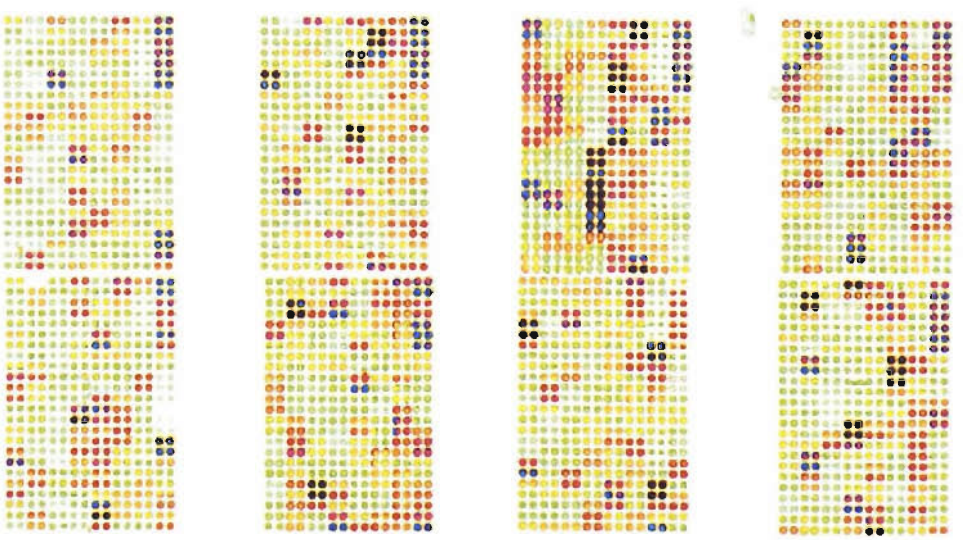


Figure 1.13. The results of a typical genotyping microarray.

1.7.1.6. Invader Assay®

Developed by Third Wave technologies the invader assay is a cleavage-based assay, which utilise the flap endonucleoase activity of cleavase enzymes. These enzymes cleave nucleic acid molecules at specific sites based on structure rather than sequence. This assay requires no PCR amplification and generates a signal only in the presence of the correct target sequence. As described in **Figure 1.14** two partially overlapping probes bind to the target DNA sequence to produce a complex recognised by the clevase enzyme. The enzyme cleaves the flap and a signal is generated by the released flap hybridising to a fret cassette which in turn releases a fluorophore.[46] A mismatch of the point of overlap blocks the cleavage of the flap. In this way discrimination is achieved. This assay has been utilised for the analysis of SNPs.[47, 48]

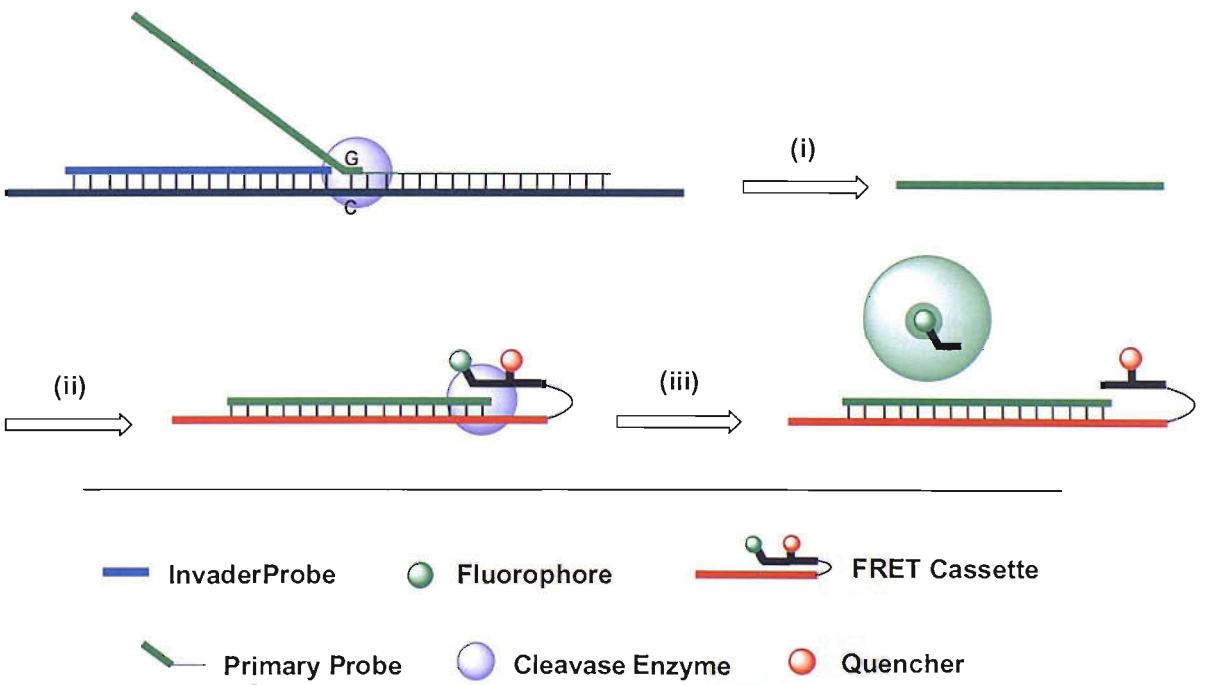


Figure 1.14. Invader® assay with fluorescence detection. (i) Cleavase enzyme releases flap. (ii) Hybridisation of cleaved flap to FRET cassette. (iii) Cleavage leads to generation of signal.

1.7.2. Direct Detection of Genetic Mutations by MALDI-TOF-MS

In order to take maximum advantage of the wealth of information for clinical or forensic purposes a wide variety of approaches for the rapid genotyping of SNPs have been developed.

Since it was first introduced in 1988 MALDI-TOF-MS has revolutionised the analysis of biomolecules. Initial work was constrained to the analysis of short thymidine containing sequences as extensive fragmentation was observed during the analysis of oligonucleotides of mixed composition. The progress of the technique was severely retarded until in 1993 Wu *et al.* discovered that 3-hydroxypicolinic acid (3-HPA) permitted the analysis of single stranded DNA in positive or negative ion mode with little fragmentation.[49] Once again DNA analysis by MALDI-TOF-MS gained momentum. The analysis of oligonucleotides by mass spectrometry has been extensively reviewed.[50-55]

One additional reason for the initially limited application of MS to the analysis of DNA is its highly anionic backbone. In solution the phosphate groups provide sites of negative charge, which form strong complexes with alkali metal cations such as sodium and potassium. This leads to the distribution of the quasi-molecular ion over a number of species of differing m/z complicating the mass spectrum and resulting in a loss in sensitivity and mass resolution. In addition, instruments with a high resolving power are required to resolve all the different species so the mass accuracy of the measurement is not affected.

Consequently, sample preparation that eliminates or minimises the formation of the metal cation adducts is crucial. A variety of quick and simple methods have been introduced for this purpose. These methods include the addition of a few beads of cation exchange resin (either proton or ammonium[56, 57]) to the oligonucleotide prior to analysis by MS or the addition of ammonium salts to the matrix.[58] The basis of this approach is to exchange the non-volatile cations with protons, which are readily lost after protonation.

The micro dialysis approach has been shown to be an effective procedure for desalting and removing low molecular weight species from oligonucleotide samples. Liu *et al.* have effectively employed an on-line dialysis method for desalting oligonucleotide of variable size and sequence prior to ES mass spectrometry.[59, 60]

1.7.2.1. Analysis of PCR Products by MALDI-TOF-MS

The ability to identify the presence of a single base substitution by a direct molecular mass measurement of a PCR product has been readily demonstrated by ES-FT-ICR-MS.[61-63] However the full capabilities of MALDI-TOF-MS for this application have not yet been realised.

In order to effectively discriminate SNPs or other substitutions by mass spectrometry the double stranded PCR amplicons have to be denatured and the analysis of the individual strands is required. Otherwise changes in mass in one strand may be neutralised by changes on the complementary strand resulting in no overall change in mass of the double stranded product. MALDI-TOF-MS has the added advantage of denaturing the duplex during the ionisation process. However the mass resolution is insufficient to distinguish between the coding and non-coding stands.

PCR products require more careful handling than synthetic oligonucleotides and the analysis is complicated by the presence of salts, enzyme, detergents and other components used in the PCR process. In addition, low molecular weight species such as primers and nucleotide triphosphates have been shown to ionise in preference to high molecular weight PCR products. For this reason the direct mass analysis of PCR amplicons requires an effective purification and desalting technique prior to analysis by MALDI-TOF-MS.

Cold ethanol precipitation from ammonium acetate has been shown to be an effective purification strategy.[64] Ethanol precipitation successfully removes enzyme, primers and nucleotide triphosphates and has the added advantage of concentrating the samples. Whilst it is a simple method in general it has to be used in conjunction with procedures that effectively remove alkali metal cations. The technique is highly molecular weight dependent and multiple cycles are often required to remove all other contaminants. Additionally, the procedure is time consuming and not amenable to automation.

In many cases PCR amplicons have been purified by capture on a solid phase. One of the PCR primers is labeled with a biotin and the resulting double stranded PCR amplicon is immobilised on a streptavidin coated magnetic bead. After washing to remove impurities the double stranded or single stranded product can be released for analysis.[65, 66]

Reverse phase column chromatography has also been demonstrated as an effective means of purifying PCR amplicons. Pipette tips filled with reverse phase high performance liquid chromatography (RP-HPLC) column media are now commercially available. The device serves as a miniature chromatography column for micro scale solid phase extraction. This approach can be easily automated and as a consequence many samples can be concentrated, purified and prepared for analysis by MALDI-TOF-MS in a short period.[67, 68]

A number of the most popular purification procedures were compared in a recent study to determine the most efficient method for purifying oligonucleotides for mass analysis.[69] The study concluded that micro-pipette tips packed with C-18 reversed phase packing material (Zip Tips) were the most effective means of purifying the smaller oligonucleotides however, microdialysis was found to be an effective alternative for purifying higher molecular weight oligonucleotides.

Direct analysis of PCR products by MALDI-TOF-MS is one of the earliest mass spectrometry based genotyping methods developed.[70, 71] Successful genotyping of short tandem repeats in a PCR product has been achieved by MALDI-TOF-MS in conjunction with solid phase purification.[72] The resolution achieved was sufficient to differentiate between any of the different nucleotides, including **A** to **T** mutations in a single stranded PCR product 69 nucleotides in length. Encouraging results were also obtained from a biplex assay. This approach was also employed for the detection of short tandem repeats.[66, 72]

1.7.2.2. Sequencing PCR Products by MALDI-TOF-MS

The MALDI ionisation technique primarily yields singly charged species, which simplifies spectral interpretation. In addition, MALDI-TOF-MS enables the recording of many data points in a single experiment, which permits the analysis of complex mixtures. As a consequence MS has readily been applied to the analysis of fragments of oligonucleotides that in turn can be used to determine the sequence of an intact oligonucleotide. Two alternative approaches have been introduced; the mass analysis of products generated from a Sanger sequencing reaction and the digestion of the template by sequence specific restriction enzymes and subsequent mass analysis of the digested products.

The Sanger sequencing method involves the generation of a set of oligonucleotides complementary to the single stranded target sequence that differ in length.[33, 73] The traditional gel electrophoresis separation is replaced with analysis by mass spectrometry. In four independent reactions a primer anneals to the template and a DNA polymerase initiates replication in the presence of a high concentration of dNTPs and a low concentration of ddNTPs. The primer is extended with ddNTPs however, extension is terminated by the incorporation of the ddNTP. As the incorporation of the ddNTP is a random event a series of oligonucleotides that differ in length are generated in each reaction mixture. The four reaction mixtures are analysed by mass spectrometry and the sequence of the DNA template can be deduced directly from the four mass spectra.[67, 74, 75]

Kirpekar and co-workers have successfully used this method for determination of a single base deletion in the human LDL receptor gene.[76] Additionally, the purine nucleotide derivatives 7-deaza-dATP and 7-deaza-dGTP were successfully used in the extension reaction as they have been shown to exhibit a greater ion stability than their natural counterparts.

Phosphodiesterase digestions are commonly used in molecular biology to sequentially remove nucleotides from a single stranded template. The combination of exonuclease digestion and mass spectrometry has been previously used for the sequencing of oligonucleotides. The exonuclease sequentially removes one nucleotide from one end of the oligonucleotide producing a ladder of degradation products. Aliquots of the digestion mixture are removed at various time intervals and spotted on the MALDI target. The fragments are analysed by MALDI-TOF-MS and the sequence of the oligonucleotide is determined by the mass differences between adjacent ions.

Piles initially demonstrated sequencing of a 12-mer oligonucleotide with the 5'-exonuclease calf spleen phosphodiesterase and the 3'- exonuclease snake venom phosphodiesterse which cleave from the 5'- and 3'-termini respectively.[58, 77, 78] This procedure has subsequently been applied to the analysis modified oligonucleotides[79] and was successfully used for the analysis of PCR products.[80]

The loss in signal intensity, mass resolution and decrease in ion abundance with increasing DNA size has proven to be the limitation to the direct analysis of PCR amplicons containing SNPs by MALDI-TOF-MS. One way to circumvent this problem is to analyse shorter amplicons that are produced in allele specific reactions.

1.7.3. Indirect Detection of Genetic Mutations by MALDI-TOF-MS

In recent years several protocols using MS have been established for the large scale scoring of SNPs and genetic variations. These methods are all based on the generation and subsequent analysis of allele specific products and have been extensively reviewed in recent publications.[81-84] The vast majority of techniques for detecting polymorphisms can be broadly classified as primer extension, enzymatic cleavage and hybridisation.

1.7.3.1. Primer Extension

Primer extension is the most commonly employed molecular biology procedure for the generation of allele specific products for SNP analysis. It has been proven to be a flexible and reliable method and in addition generates small products for detection by mass spectrometry.

Several different variations in assay design and reaction conditions have been used, however all assays employ a thermo-stable DNA polymerase to distinguish between the two alleles by addition of a nucleotide or series of nucleotides.

Following an initial PCR step, excess primers and dNTPs are removed by capture on a magnetic bead or by treatment with shrimp alkaline phosphatase (SAP) and exonuclease I. A second set of primers anneal directly adjacent to the SNP site and the allele specific products are generated using solely ddNTPs or a specific combination of ddNTPs and dNTPs. The extension specifically incorporates nucleotides that are complementary to the DNA template and terminates at the polymorphic site or upstream of the site with the inclusion of a ddNTP.

Following a further purification step the extension products are analysed by MALDI-TOF mass spectrometry. The composition of the variable site is determined from the mass of the extension product detected.

PRimer Oligonucleotide Base Extension Assay

The primer oligonucleotide base extension or PROBE assay was first introduced in 1997 by Braun and co-workers.[65, 85-87] The template DNA is first amplified by PCR in the presence of one 5'-biotinylated primer and one unmodified primer.

As described in **Figure 1.15**, the PCR amplicons are captured on a streptavidin coated magnetic bead and washed to remove excess primers and dNTPs. Following denaturation, the immobilised single strand serves as a template for a primer extension reaction. A second primer anneals adjacent to the variant site. In the presence of three dNTPs and one ddNTP the 3'-terminus of the primer is extended with extension terminating by the incorporation of the ddNTP.

Typically one allele is extended by one base and the second allele is extended by two or three bases generating a mass difference of approximately 300 Da between extension products. The extension products are denatured, desalted and concentrated by ethanol precipitation. Determination of the mass of the extended primer identifies the base at the polymorphic site in the template DNA.

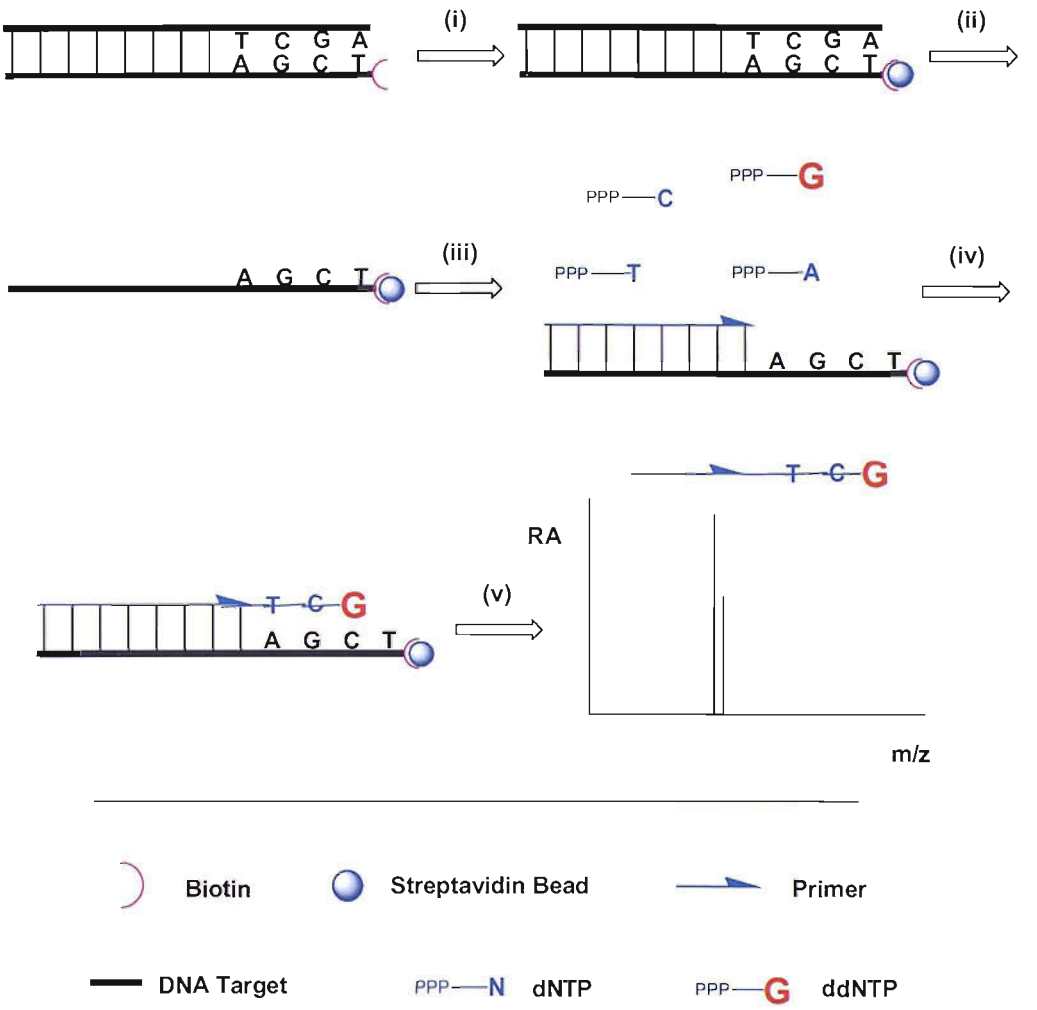


Figure 1.15. The PROBE assay. (i) A double stranded biotinylated PCR product is immobilised on a streptavidin bead. (ii) The duplex is denatured and the unbound strand is washed away. (iii) A primer anneals upstream of the varient site. (iv) The 3'-terminus of the primer is extended by a DNA polymerase in the presence of three dNTPs and one ddNTP. The extension is terminated by the incorporation of the ddNTP. (v) The extension product is analysed by MALDI-TOF-MS.

This method was further developed and optimised by Sequnom (San Diego, CA) and marketed as the massEXTEND assay governed by the massARRAY system.[83, 88, 89] The massARRAY system offers several advantages; namely it is a single tube reaction, which employs an enzyme to destroy the remaining PCR primers and remaining dNTPs prior to primer extension reaction. Following primer extension on both strands of the PCR amplicon the primer extension products are desalted by addition of NH_4^+ cation exchange beads and analysed by MALDI-TOF-MS. Since this procedure requires only liquid handling steps it is readily compatible with automated liquid handling systems and thermo-cycling systems. One additional attraction of the massEXTEND platform is the use of a silicon chip MALDI target preloaded with the 3-HPA matrix. The sample is loaded onto a spot of similar size to the area of laser irradiation during ionisation. This enables the maximum quantity of sample to be assessed in a single laser shot. Additionally, the spot is more homogenous which substantially reduces the need for searching for 'sweet' spots or parts with a higher concentration of analyte. Importantly, the increased shot-to-shot and sample-to-sample reproducibility makes it possible to perform high-throughput and quantitative analysis.

Li *et al.* have utilised this approach in an assay composed of three steps; PCR amplification, phosphatase digestion and single base primer extension.[90] A modified primer containing a 5'-biotin and a cleavable nucleotide at the 3'-terminus is used in the primer extension. Following capture on the solid phase the nucleotide on the 3'-end of the primer is released and detected by MALDI-TOF-MS. In the case of the heterozygote the mass difference between the two extension products is used to determine the nucleotide present at the SNP site. This method has many advantages, namely the reduced size of the allele specific product detected directly leads to a higher sensitivity and resolution in MALDI-TOF detection.

Pinpoint Assay

A similar approach has been employed in the pinpoint assay, which was later commercialised by Applied Biosystems as the Sequazyme typing assay (**Figure 1.16**).[68, 91] After PCR amplification of the region of interest the residual dNTPs and primers are degraded by SAP and exonuclease I. A primer extension reaction is carried out with four ddNTPs resulting in the addition of a single base. The primer extension products are then purified and concentrated using Zip Tips or by gel filtration. Using the unextended primer as an internal calibrant, the ions observed in the mass spectrum not only reflect the mass of the particular primer extension product, but also identify the added nucleotide thus revealing the nature of the base at the polymorphic site.

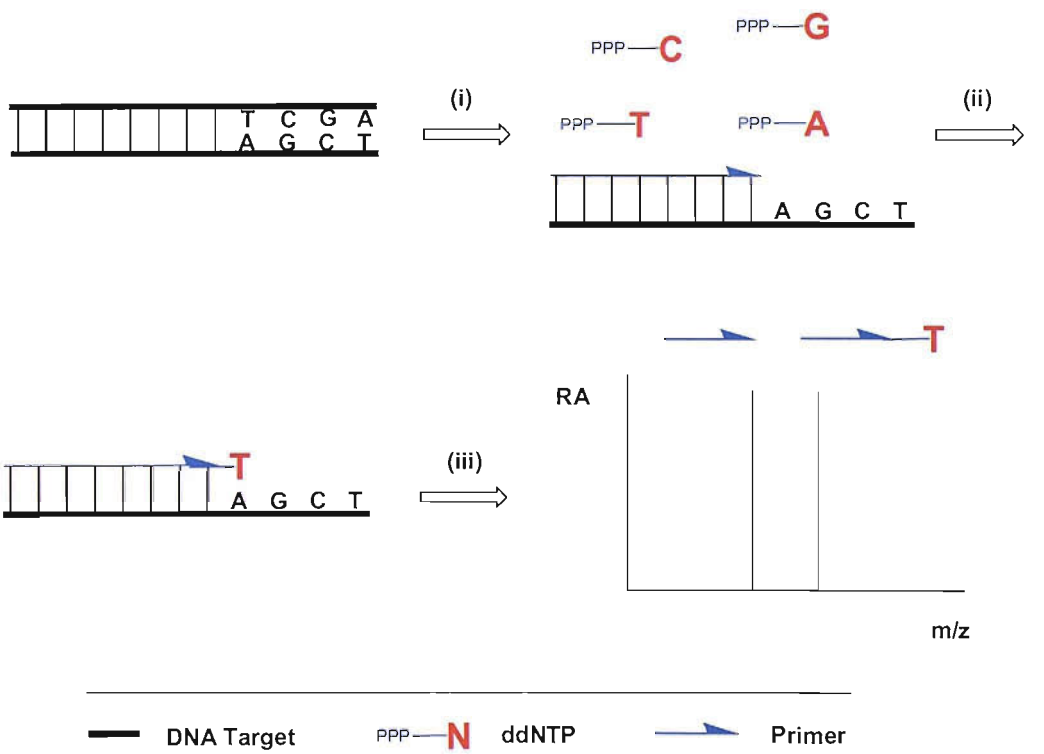


Figure 1.16. The pinpoint assay. (i) Following PCR a primer anneals upstream of the varient site. (ii) The 3'-end of the primer is extended by a DNA polymerase in the presence of four ddNTPs. A single base is added complementary to the variant base. (iii) The extension product is analysed by MALDI-TOF-MS.

Multiplex assays have successfully been carried out using primers of successively increasing lengths designed to avoid mass interference between primers and extension products.[92-94] In order to avoid overlap of the primers and extension products and to prevent the generation of high molecular weight products, the level of multiplexing is constrained to 12 to 15 polymorphic sites.

One limitation of this approach stems from the small mass difference between some nucleotides namely; **A** and **T** (9 Da). As a consequence it is difficult to unambiguously determine the base added on low specification mass spectrometers. This problem has been overcome by the use of mass tagged primers. Whilst the 3'-portion of the primer is complementary to the target, a mass tag composed of a variable number of thymidylic acid units[95] is attached to the 5'-terminus. Within a given mass range, the number of typed loci can be increased avoiding potential overlaps between primers and extended primers.

Using this approach, the need for extensive sample clean up and a mass spectrometer with a high resolving power is alleviated. However, due to instrumental limitations in terms of resolution and sensitivity, the size of the extension products is still restricted to 40 bases. An additional variation of this method was reported by Fei,[96] Kim[97] and Abdi[98] who have utilised modified ddNTPs in the primer extension reaction to further increase the mass difference between variants.

Whilst the pinpoint assay has been applied to several multiplex studies the majority of them have avoided the analysis of **A**/**T** mutations. In addition, the high levels of multiplexing are only possible with the time consuming and cumbersome optimisation of the multiplex PCR and primer extension reactions.

VSET Assay

The very short extension or VSET assay is an alternative approach introduced by Sun and co-workers.[99, 100] Following PCR amplification of specific fragments of the DNA template, primer extension is carried out in the presence of a specific combination of three ddNTPs and one dNTP. This gives rise to the extension of one allele by one base and the other by two bases. The primer extension products are purified by ethanol precipitation and analysed by MALDI-TOF-MS. This method addresses some of the problems associated with the pinpoint and PROBE formats. Compared with the PROBE assay the VSET assay produces shorter extension products. This results in increased detection limits, as the primer extension reaction is more efficient and shorter DNA products are more readily analysed by MALDI-TOF-MS than longer oligonucleotides. In contrast to the pinpoint format where the desalting and purification step is essential to clearly differentiate between allele specific products, stringent purification is not essential as the base at the variant site is determined by the number of nucleotides added. In addition, the multiplexing potential of the assay have been evaluated.[100] The mass differences in primer extension products were enhanced by addition of a variable number of **T** units to the 5'-region of the primer extension primer. In this way this method has the potential to analyse 15 SNP sites simultaneously in a single assay.

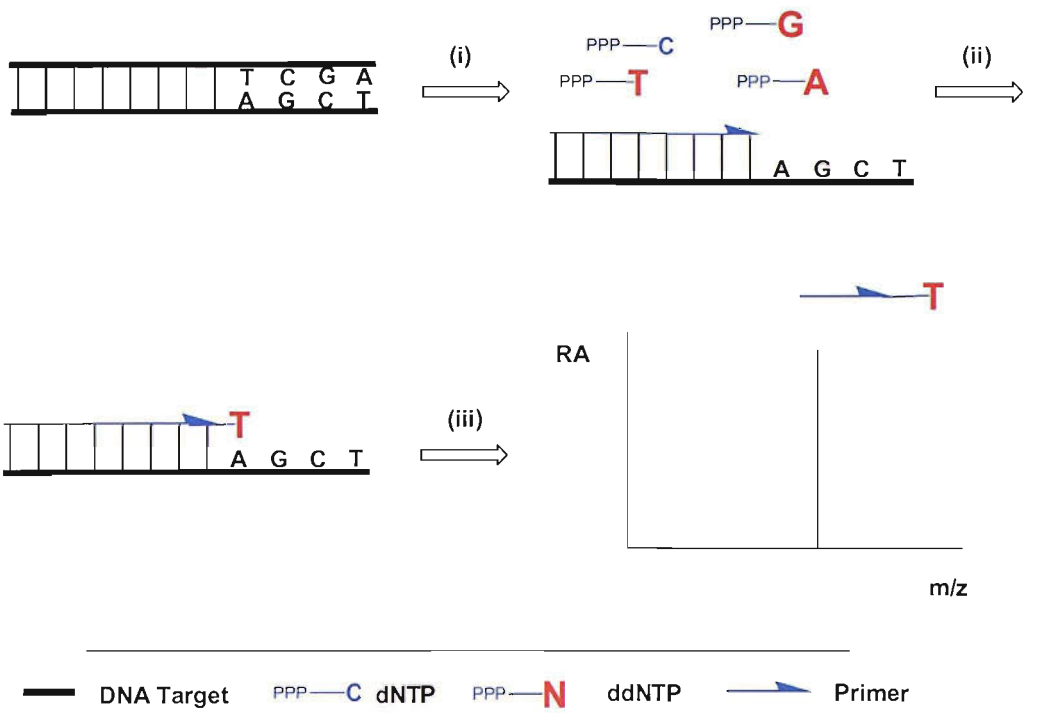


Figure 1.17. The VSET assay. (i) Following PCR a primer anneals upstream of the varient site. (ii) The 3'-end of the primer is extended by a DNA polymerase in the presence of three ddNTPs and one dNTP. (iii) The extension product is analysed by MALDI-TOF-MS.

GOOD Assay

First introduced by Sauer *et al.* in 2000, the GOOD assay has since been developed into a high throughput method of genotyping SNPs.[101, 102] As shown in **Figure 1.18**, after PCR amplification of the region of template DNA containing the variant site, residual dNTPs are degraded with SAP. Primer extension is then carried out in the presence of α-S-ddNTPs using an allele specific primer in which the 3'-region is prepared using α-S-dNTPs. The unmodified sections of the primer extension products are removed by 5'-phosphodiesterase digestion. The first phosphothioate group at the 5'-terminus inhibits digestion leaving only the core 4 - 5 modified nucleotides that contain the allele specific information.

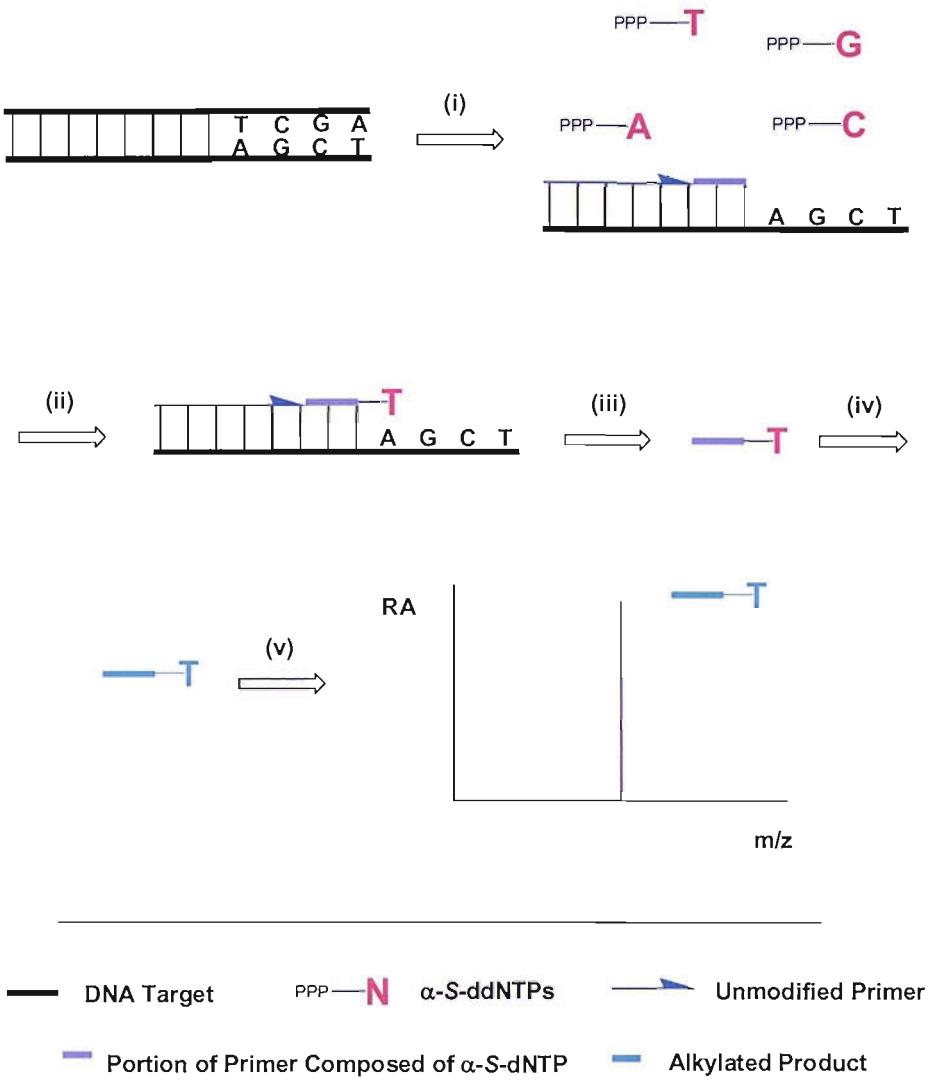


Figure 1.18. The GOOD assay. (i) Annealing of the modified primer to the PCR amplicons. (ii) The 3'-end of the primer is extended by a DNA polymerase using α-*S*-ddNTPs. (iii) Phosphodiesterase digestion. (iv) Alkylation of products with methyl iodide. (v) The alkylated product is analysed by MALDI-TOF-MS.

Chemical modification of the phosphothioate groups with methyl iodide neutralises all charges on the backbone, which enables the modification of the product for analysis by positive or negative ion MALDI-TOF-MS. Charge tagging *via* introduction of a single positive charge on an amino modified base[103] and backbone neutralisation[104] has proven to leave the products less susceptible to cation adduct formation and insensitive to contaminants in the reaction buffers. This enables the assay to be carried out in a single tube, using only incubation and thermal cycling steps.

An alternative approach was employed using an extension primer containing phosphothioates with a photo-cleavable linker separating the modified and unmodified regions.[105] The phosphodiesterase digestion is omitted as the unmodified part of the primer is cleaved by exposure to UV light. Following alkylation the negatively charge tagged product was successfully detected by MALDI-TOF-MS. The assay was subsequently modified and streamlined into three simple steps; PCR amplification, primer extension and phosphodiesterase digestion avoiding the use of methyl iodide.[106] This was achieved by the use of primers prepared using methyl phosphonates and a novel DNA polymerase in the primer extension reaction. The polymerase recognises methylphosphonate containing primers and preferentially incorporates ddNTPs over dNTPs alleviating the SAP digestion and alkylation steps.

In summary, all variations on the GOOD assay are based on the charge tag concept where the allele specific product is conditioned to carry either a single positive or negative charge. The increase in detection sensitivity observed when using the modified products over unmodified products enables the analysis of the crude reaction mixture alleviating the need for purification. The entire assay can be carried out in a single tube and can be carried out using liquid handling systems rendering it amenable to high through-put applications.

1.7.3.2. Cleavage

Invader Assay®

In contrast to the other methods described here the invader assay® does not require an initial PCR amplification step.[107] As described previously the assay involves the sequence specific hybridisation of two oligonucleotide probes to form an overlapping structure at the variant site. The probe oligomer anneals upstream of the variant site is composed of two sections. The 3'-region is complementary to the target DNA whilst the 5'-section acts as the allele specific signal oligonucleotide. The oligomer which binds downstream of the variant site is termed the Invader probe. The reaction is carried out in a single tube at the melting temperature of the Invader probe and template duplex. Under these conditions the invader probe momentarily occupies the complementary site on the DNA template forming an overlap structure. A flap endonuclease (FEN) specifically recognises the structure and cleaves the unpaired region at the 5'-region of the probe oligomer resulting in a 3'-hydroxyl terminating cleavage product. This results in a linear accumulation of the cleavage product or signal oligomer. The selectivity of the assay arises from the formation of the correct configuration of the Invader and probe. If the incorrect invasive structure is created, for example, the wild type probe and mutant DNA target cleavage is prohibited.

Subsequently the assay was modified and re-introduced as the invader squared assay designed for fluorescence[108, 109] detection in addition to detection by MS.[110, 111] Different probes are used in two sequential stages of the reaction. As described in **Figure 1.19** the primary invader reaction is carried out resulting in accumulation of the first cleavage product, which in turn, is designed to function as an invader probe in the secondary reaction. A second probe oligomer is then introduced.

The cleavage product of the second probe is a biotin labeled low molecular mass signal molecule, which is isolated by capture on a streptavidin coated magnetic bead and subsequently analysed by MALDI-TOF-MS. The ion observed in the mass spectrum reveals the identity of the signal molecule, which in turn can be related to the primary cleavage product and genotype of the DNA template. In this way an exponential increase in the generation of the signal molecule is observed, alleviating the need for an initial PCR stage. Eliminating the PCR amplification is desirable due to the limitations of PCR in a high-throughput environment namely; contamination issues and variation in reaction conditions and yields.[112] The multiplexing capability of the assay has been demonstrated by the simultaneous analysis of twelve SNPs directly from genomic DNA.[111] The assay has been proven to be a simple and robust procedure, however it is still limited by the high cost of the flap endonuclease.

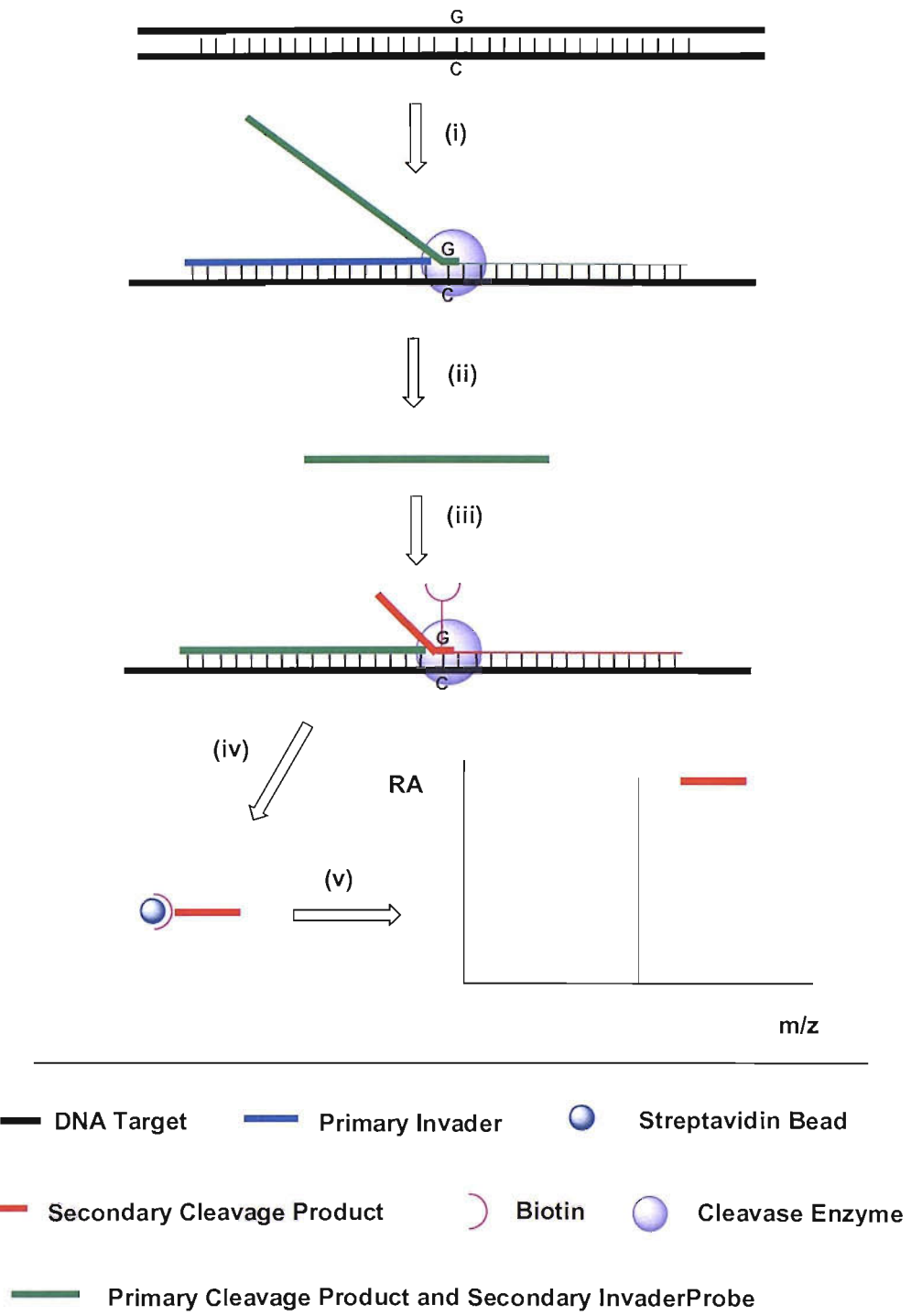


Figure 1.19. The invader assay®. (i) The primary invader and probe anneal to the DNA template. (ii) The 5'-region of the probe is cleaved by the endonuclease. (iii) The cleavage product acts as a invader for the second reaction. (iv) The unpaired region at the 5'-end of the probe is cleaved by a flap endonuclease and immobilised on a streptavidin bead. (v) Analysis by MALDI-TOF-MS.

1.7.3.3. Hybridisation

Allele specific hybridisation can be used for mutation detection and SNP analysis. Two probes, each representing one allele of the target sequence are prepared. Under the assay conditions, only the probe that is exactly complementary to the target sequence will anneal. Any single base mismatches will alter the thermodynamics of binding so that the probe will not bind. The probes are generally short and designed so that the SNP site is centrally placed so any mismatches will have the maximum disruptive effect on the duplex. Hybridisation is a physically controlled process and for this reason hybridisation based assays lack the discrimination of assays, which employ an enzyme. Hybridised (fully complementary) and unhybridised probes (probes which contain a singe base mismatch) must be separated prior to analysis by mass spectrometry. The genotype of the sample is determined by the mass of the detected probe.

Peptide Nucleic Acid Hybridisation Assay

This approach has been utilised by Ross,[113] Jiang Baucom[114] and Griffin[115] who have explored the use of peptide nucleic acid probes (PNA) in a hybridisation assay coupled with MALDI-TOF-MS for the detection of sequence polymorphisms in PCR amplified genomic and mitochondrial DNA. PCR of the target sequence is performed in the presence of one biotinylated and one unmodified primer. As described in **Figure 1.20**, the PCR product is immobilised on a streptavidin magnetic bead. The duplex is denatured and the unbound strand is washed away. PNA probes representing the two possible alleles are added and the fully complementary probe hybridises to the DNA target. Stringent washing removes any uncomplementary probes and the PNA/DNA/bead complex is then analysed by MALDI-TOF-MS.

Whilst the target DNA remains immobilised on the MALDI target the PNA probe is denatured from the DNA during the application of the matrix. The sequence of the target DNA is determined by the mass of the probe detected by MALDI-TOF-MS.

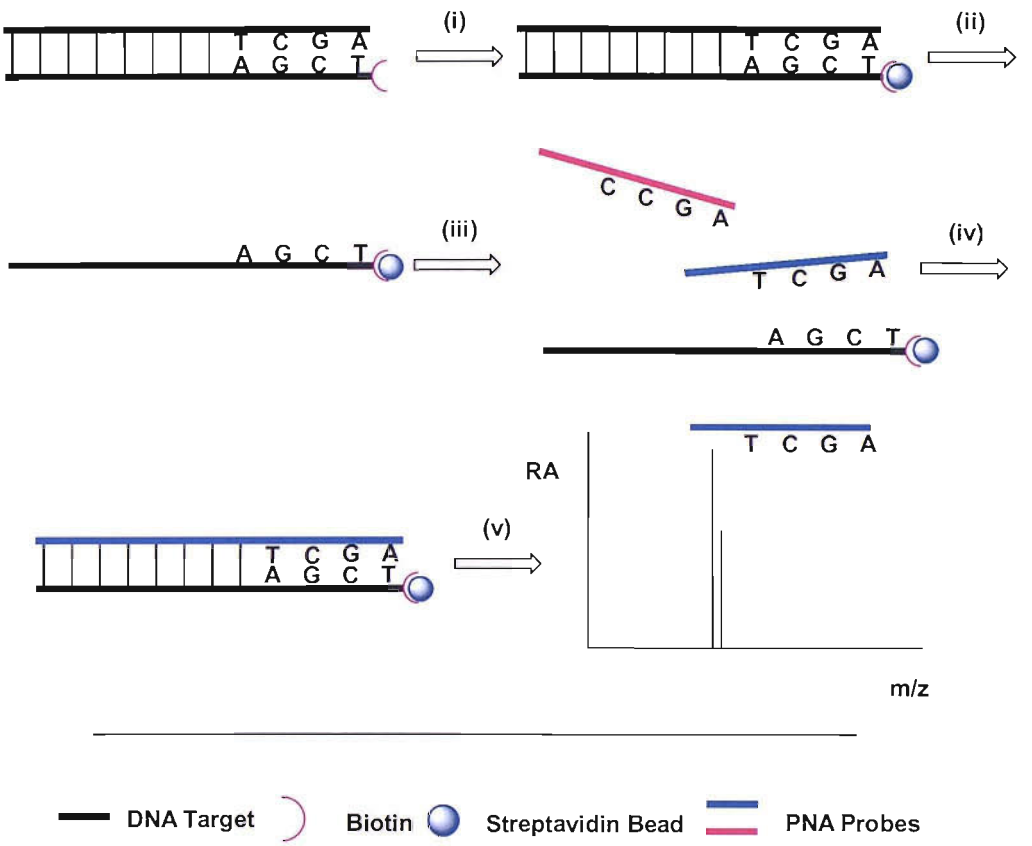


Figure 1.20. PNA hybridisation assay. (i) A double stranded biotinylated PCR product is immobilised on a streptavidin bead. (ii) The duplex is denatured and the unbound strand is washed away. (iii) PNA probes complementary to the two possibe alleles are added. (iv) Uncomplementry probes are removed by heating and washing. (v) The PNA/DNA/bead complex is analysed by MALDI-TOF-MS.

Griffin *et al.* have successfully coupled a variable number of 8-amino-3,6,dioxaoctanoic acid units to the amino terminus of the PNA probes to enhance the mass difference between similar mass probes. These modified probes have effectively been used for determination of tyrosine exon 4 polymorphisms (**Figure 1.21**).

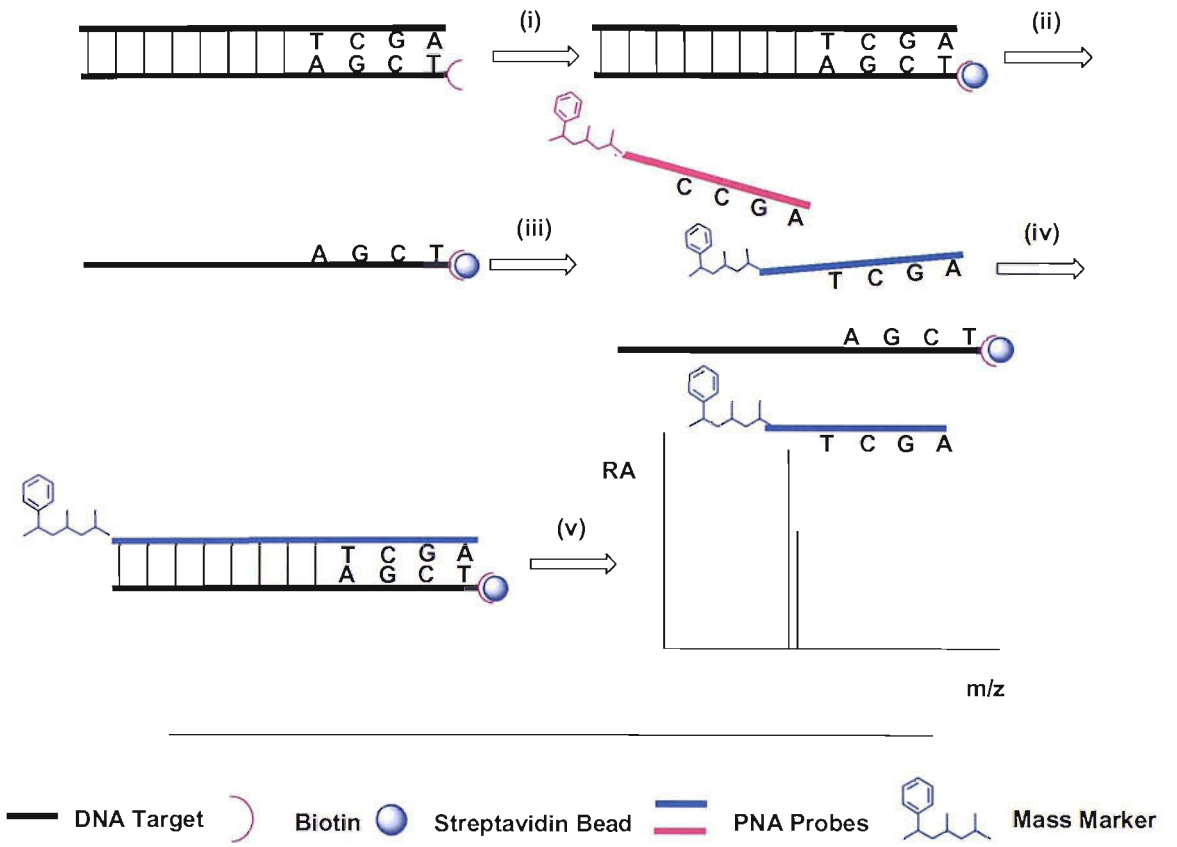


Figure 1.21. PNA hybridisation assay. (i) A double stranded biotinylated PCR product is immobilised on a streptavidin bead. (ii) The duplex is denatured and the unbound strand is washed away. (iii) PNA probes complementary to the two possibe alleles are added. (iv) Uncomplementry probes are removed by heating and washing. (v) The PNA/DNA/bead complex is analysed by MALDI-TOF-MS.

Whilst a wide variety of assays and many methods have been introduced and new techniques are still emerging, no single method is suitable for all genotyping studies and has met all the requirements in terms of throughput and cost. In light of this, a major analytical challenge is to provide the technologies for a high-throughput and cost effective method for SNP genotyping. Few of the current SNP genotyping techniques can cope with the resulting demands concerning sample throughput, automation, accuracy and cost-effectiveness.

In the work presented here a PNA hybridisation approach has been developed using sequence specific PNA probes modified with photo-cleavable mass markers for detection of point mutations. These probes have been modified at the amino terminus with a photo-cleavable linker, lysine and a low molecular weight amino acid. In addition, an isotopically labelled Peptide Organic Acid (POA) monomer is incorporated to provide each sequence with a cleavable mass marker with a unique molecular mass.

In the hybridisation assay, PNA probes representing the two alleles are added to a region of single stranded PCR amplified DNA immobilised on a magnetic bead (as described in **Figure 1.22**). The fully complementary probes hybridise to the DNA target and the uncomplementary probes are washed away. The PNA/DNA/bead complex is then analysed by MALDI-TOF-MS. The genotype of the sample is determined by the quasi-molecular ion of the mass marker released photo-chemically during laser ablation.

The probes used in this assay are primarily composed of PNA. In addition to showing excellent discrimination for mismatches, PNA is significantly easier to analyse and detect by mass spectrometry over native DNA. As a consequence this type of assay does not require extensive purification or chemical modification of the probes prior to analysis by mass spectrometry.

This procedure differs from the existing PNA hybridisation assays since it utilises a cleavable low molecular weight mass marker. The mass markers give each PNA probe a unique molecular mass identifier and maintains the molecular mass within the 500 – 1000 Da mass range. This permits the use of unique isotopic labels and enables the use of a proprietary software package, which will automatically identify the designated patterns. This offers the potential for development into a high-throughput, powerful screening method for detection of SNPs and point mutations.

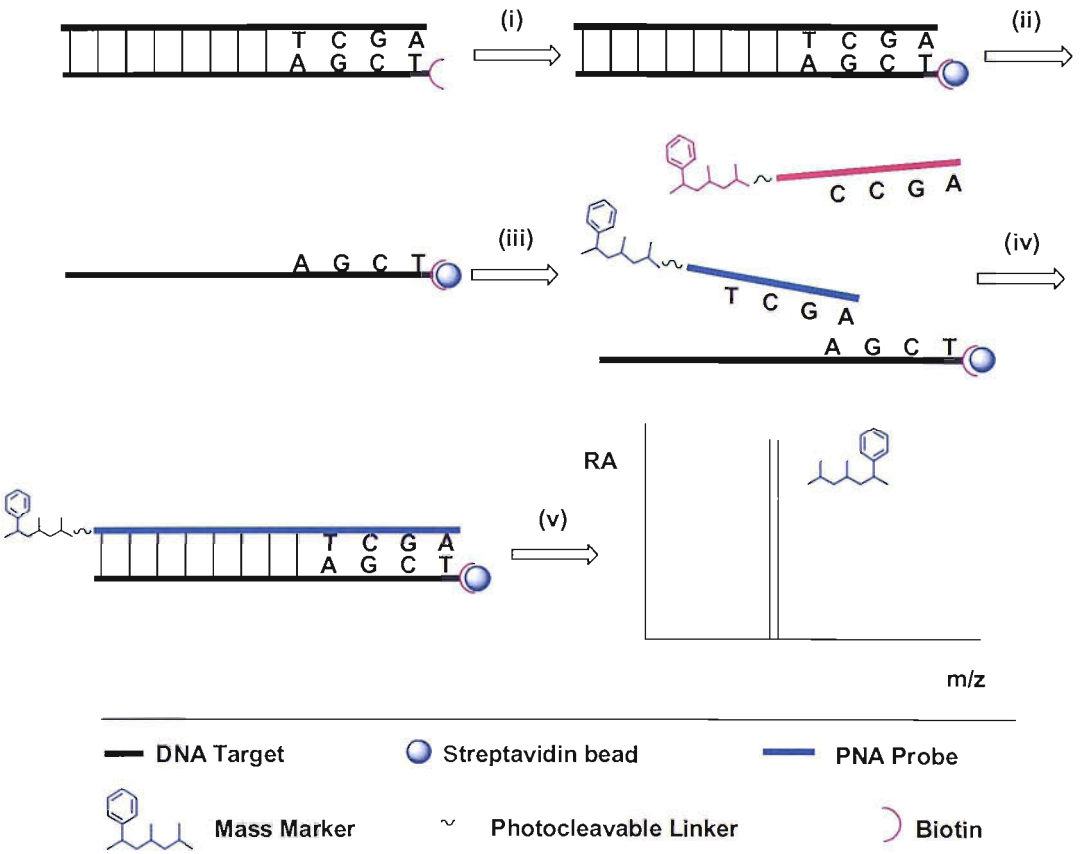


Figure 1.22. PNA hybridisation assay. (i) A double stranded biotinylated PCR product is immobilised on a streptavidin bead. (ii) The duplex is denatured and the unbound strand is washed away. (iii) Modified PNA probes complementary to the two possibe alleles are added. (iv) Uncomplementry probes are removed by heating and washing. (v) The PNA/DNA/bead complex is analysed by MALDI-TOF-MS. The genotype of the sample is determined by the quasi-molecular ion of the mass marker cleaved from the PNA probe.

2. Preliminary Work and Pilot Study

2.1. Preliminary Work

2.1.1. Peptide Nucleic Acids

As shown in **Figure 2.1**, PNA is a structural analogue of DNA developed by Nielsen *et al.* in which the sugar phosphate backbone has been replaced with an uncharged, achiral polyamide chain.[116, 117] As the spacing between the natural nucleobases is retained, PNA oligomers are able to form the same Watson Crick base pairs as regular oligonucleotides. This enables PNAs to form duplexes and triplexes with the complementary DNA or RNA sequences.[118, 119]

PNA oligomers can have amino- and carboxy- ends analogous to the 5'- and 3'- termini of DNA, and readily form duplexes with the complementary DNA strand in either orientation. However, the antiparallel orientation is strongly preferred. The heteroduplexes formed between PNA and DNA exhibit a higher thermal stability than the equivalent DNA homoduplexes. The formation of a PNA/DNA duplex can also proceed in the absence of Mg^{2+}, a factor that destabilises DNA/DNA duplexes. In contrast to DNA/DNA duplexes, the hybridisation of a PNA/DNA duplex is virtually independent of salt concentration. Adjusting the ionic strength of the buffer can be a very useful parameter when designing procedures where competing DNA is present in the sample.[120-122]

All of these properties are a result of the lack of electrostatic repulsion between negatively charged phosphate groups, which is encountered when complementary oligonucleotides hybridise. Additionally, PNA oligomers have shown superior sequence discrimination. Studies have shown that a single base mismatch in a PNA/DNA is less stable than the same mismatch in the corresponding DNA/DNA duplex.[121, 122]

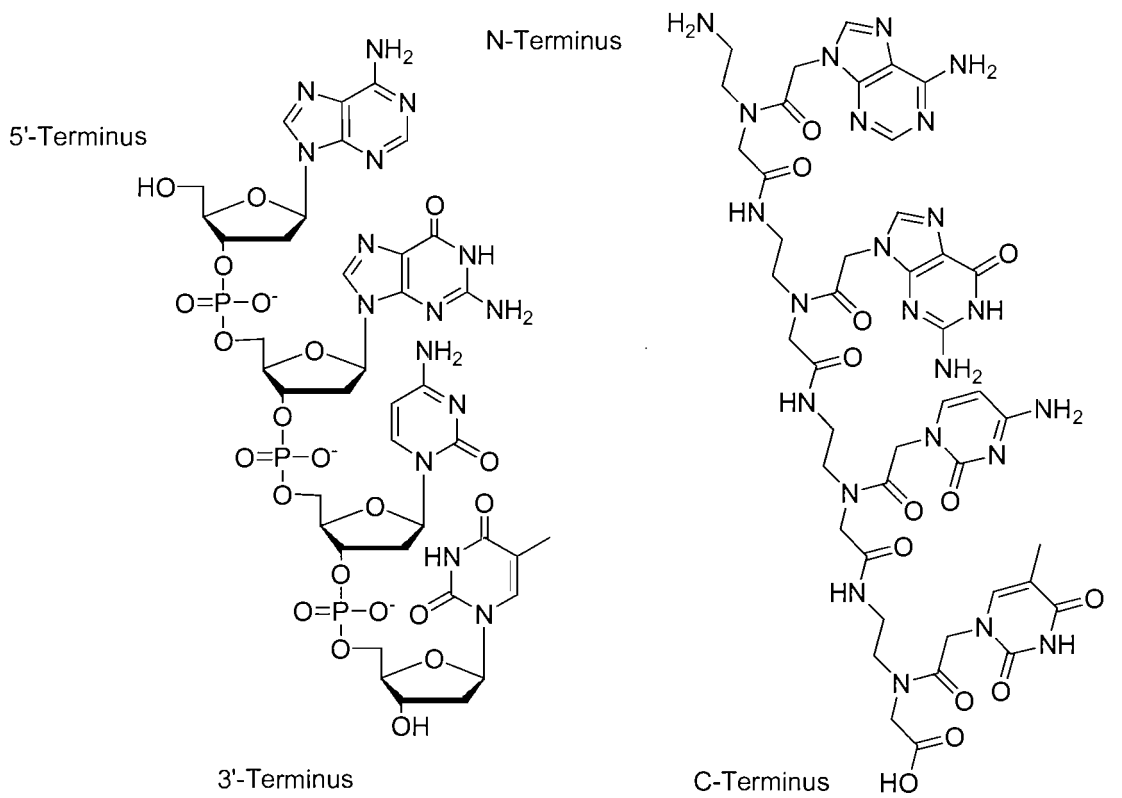


Figure 2.1. The chemical structures of DNA (left) and PNA (right).

The modification to the backbone means that PNA oligomers are not recognised or degraded by enzymes such as proteases, nucleases and polymerases. For this reason PNAs cannot be copied or used directly as PCR primers or incorporated directly into PCR amplicons. However, this high bio-stability is important for their use as antisense and antigene therapies.[123] This also provides assays with a better shelf life and makes PNAs an ideal candidate for use in diagnostic and molecular biology applications.[124-126]

The preparation of PNA oligomers from individual monomers is based on standard solid phase peptide synthesis procedures. Oligomer synthesis and the preparation of the individual PNA monomers are discussed in detail in **Chapter 4**.

PNAs have been extensively studied by mass spectrometry. Griffith has successfully detected a PNA/DNA duplex by ES-MS.[127] This method was subsequently utilised by Sforza *et al.* to study the orientation of the heteroduplexes formed between DNA and PNAs modified with D-lysine (known as chiral box PNAs).[128] Achiral and chiral box PNA/DNA duplexes were studied in negative ion mode. Analysis by mass spectrometry demonstrated the formation of a stable parallel and antiparallel duplex between the achiral PNA and DNA targets. In contrast, the chiral box PNA did not form a duplex with the parallel DNA target.

Flora and co-workers have examined the gas phase fragmentation pathways of PNAs by ES-FT-ICR-MS.[129-131] Complete sequence information was obtained from short mono-deprotonated PNA oligomers by MS-MS, in this case sustained off-resonance irradiation-collision induced dissociation (SORI-CID). The studies showed the gas phase fragmentation behaviour of PNAs resembles that of both peptides and oligonuclotides. A nomenclature for the product ions observed on the fragmentation of PNA oligomers was proposed incorporating aspects of peptide and oligonucleotide fragmentation.

Butler *et al.* have studied PNAs by MALDI-TOF mass spectrometry.[132] PNA oligomers are more efficiently analysed in the matrices used to study protein or peptides than conventional oligonucleotide matrices because of the peptide backbone. PNAs survive the MALDI process intact forming predominantly singly charged species, in contrast to oligonucleotides where fragmentation due to depurination is often observed.

2.2. Pilot Study

2.2.1. Cystic Fibrosis

Cystic fibrosis (CF) is the most common potentially lethal autosomal recessive disease in caucasians. In the UK, CF affects over 7,500 people. 2.3 million or 1 in 25 of the population carry faulty CF genes. Cystic fibrosis shows a range of forms from mild to severe. The epithelial cells lining a patients internal organs produce excessive sticky mucus that clogs the lungs and leads to life-threatening lung infections. These thick secretions also obstruct the pancreas, preventing digestive enzymes from reaching the intestines to help break down and absorb food. Patients with CF have a variety of symptoms, which include; very salty-tasting skin, persistent coughing, wheezing or shortness of breath and an excessive appetite with poor weight gain.

The abnormal gene that causes CF is found on chromosome number 7 (ABCC7 gene), which encodes the cystic fibrosis transmembrane conductance regulator (CFTR) protein. The CFTR gene sequence has been identified and a number of mutations have been found which give symptoms that range in severity. Five loci, in particular, account for the mutations that cause 90 % of the disease. W1282X (**G** to **A** mutation), N1303K (**C** to **G** mutation), G542X (**G** to **T** mutation) and G551D (**G** to **A** mutation) are single point mutations which occur at different sites whilst ΔF508 is a three base (**CTT**) deletion.[133, 134]

2.2.2. PNA Probes

Whilst the use of oligonucleotides in mutation detection has distinct advantages, their detection by MS is limited by the significant loss in signal intensity observed in the mass spectrum and mass resolution with increasing DNA size. As discussed in **Chapter 1**, one of the contributing factors is the tendency of the polyanionic backbone to form strong complexes with cations. The neutral PNA backbone, does not readily form cation adducts which hamper sensitivity and resolution, can be seen in **Figure 2.2**. Furthermore, the analysis of DNA by mass spectrometry is sensitive to contaminants such as detergents, buffers, amplification enzyme and other components used in PCR reactions. Whilst many protocols exist for the purification of DNA products, these are difficult to automate and increase the cost per analysis. PNA oligomers have been shown to exhibit superior hybridisation properties and are readily analysed by mass spectrometry (**Figure 2.2** depicts mass spectra of 10 pmol of DNA and 100 fmol of PNA). For these reasons PNA probes were chosen for the hybridisation assay over conventional DNA probes.

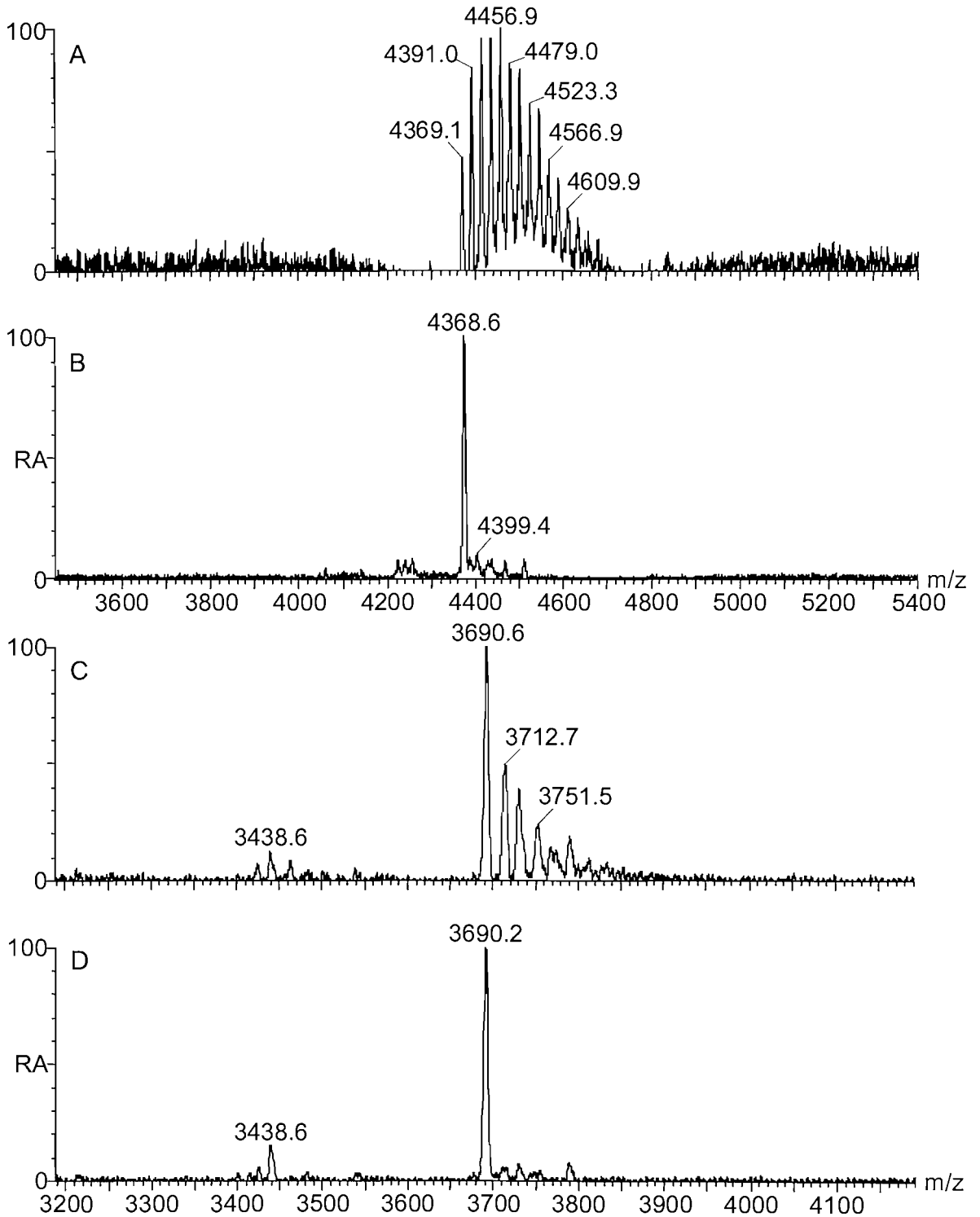
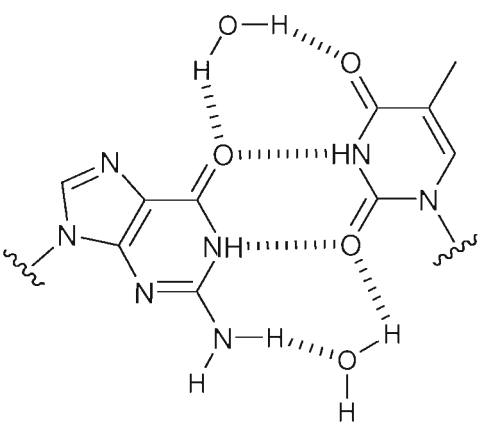


Figure 2.2. Linear positive ion mass spectrum. (A) 10 pmol of DNA (ACAGTGAAGGAAAG) without H^+ dowex beads. (B) 10 pmol of DNA (ACAGTGAAGGAAAG) with H^+ dowex beads. Expected mass $(M+H)^+$ 4368.8 Da. (C) 100 fmol of PNA (PNA(14C)) without H^+ dowex beads. (D) 100 fmol of PNA (PNA(14C)) with H^+ dowex beads. Expected mass $(M+H)^+$ 3690.6 Da.

The interaction between PNA and its complementary DNA leads to the formation of more thermally stable duplexes than the natural nucleic acids. As a consequence of its enhanced stability, greater discrimination between fully complementary and mismatched duplexes is observed with a PNA/DNA duplex compared with the corresponding DNA/DNA duplex. For this reason, in general, shorter PNA oligomers are used in applications (typically 12 to 17 units) compared with the conventional DNA probes (20 – 25 nucleotides).[135] Short probes offer higher specificity since the presence of a single base mismatch has a higher destabilising impact in a short duplex than in a longer duplex, where there are adjacent regions of duplex to stabilise the mismatch. For the initial studies PNA probes 14 bases in length were chosen. The stability of a single base mismatch in a PNA/DNA duplex depends both on its location and the chemical nature of the mismatch. To have the maximum disruptive effect the variant base was placed in the centre of the probe.

The cystic fibrosis W1282X mutation involves the substitution of a **G** with an **A**. The study of a **G** to **A** mutation would involve the preparation of allele specific PNA probes with either a **C** (wild type) or a **T** (mutant) at the variant site. As the selectivity of a hybridisation assay arises from the hybridisation step, it is essential that any single base mismatches will sufficiently change the thermodynamics of binding so that under the assay conditions only the probe that is exactly complementary to the target sequence will anneal. Studies have shown that the most stable mismatched pairs in a PNA/DNA duplex are **G**:**T**.[136] The **G** and **T** bases pair through two imino-carbonyl hydrogen bonds which produces little distortion of the phosphodiester backbone and requires no protonation at pH 7 as shown in **Figure 2.3**.[137]

Additionally, water molecules cluster around the base pair, known as a wobble base pair, taking up all of the hydrogen bonding capabilities of the bases. These bridging water molecules serve to stabilise the base-base interaction. Instead of the correct **G**:**C** base pair, a **G**:**T** mismatch would be formed between the wild type DNA target (**G**) and mutant PNA probe (**T**). This would make it difficult to readily distinguish between the wild type and mutant DNA targets. To circumvent this problem PNA probes were prepared to study a **G** to **T** transversion mutation on the W1282X locus (**Figure 2.4**) *i.e.* PNA probes with a **C** (wild type) and **A** (mutant) at the variant site. During the preliminary stages this would simplify the assay and allow the evaluation of the assay conditions and mass spectrometry parameters without any additional complications.



G : T

Figure 2.3. The wobble base pair formed in a G:T mismatch.

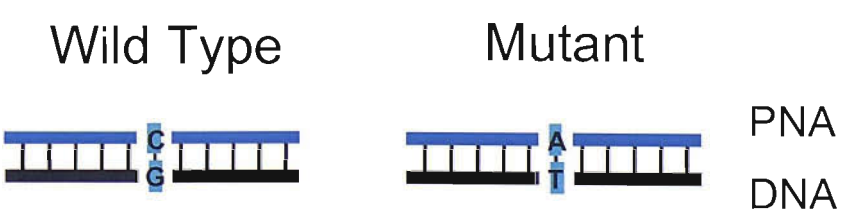


Figure 2.4. G to T transversion mutation.

2.2.3. Design of the Mass Markers

In order to successfully determine the benefits of using PNA probes with photo-cleavable mass markers in mutation detection and SNP analysis, it was necessary to prepare the three following different types of synthetic PNA probes (**Figure 2.5**)

- Unmodified PNA probe (**A**).
- PNA probe modified with a mass marker (**B**).
- PNA probe modified with a photo-cleavable mass marker (**C**).

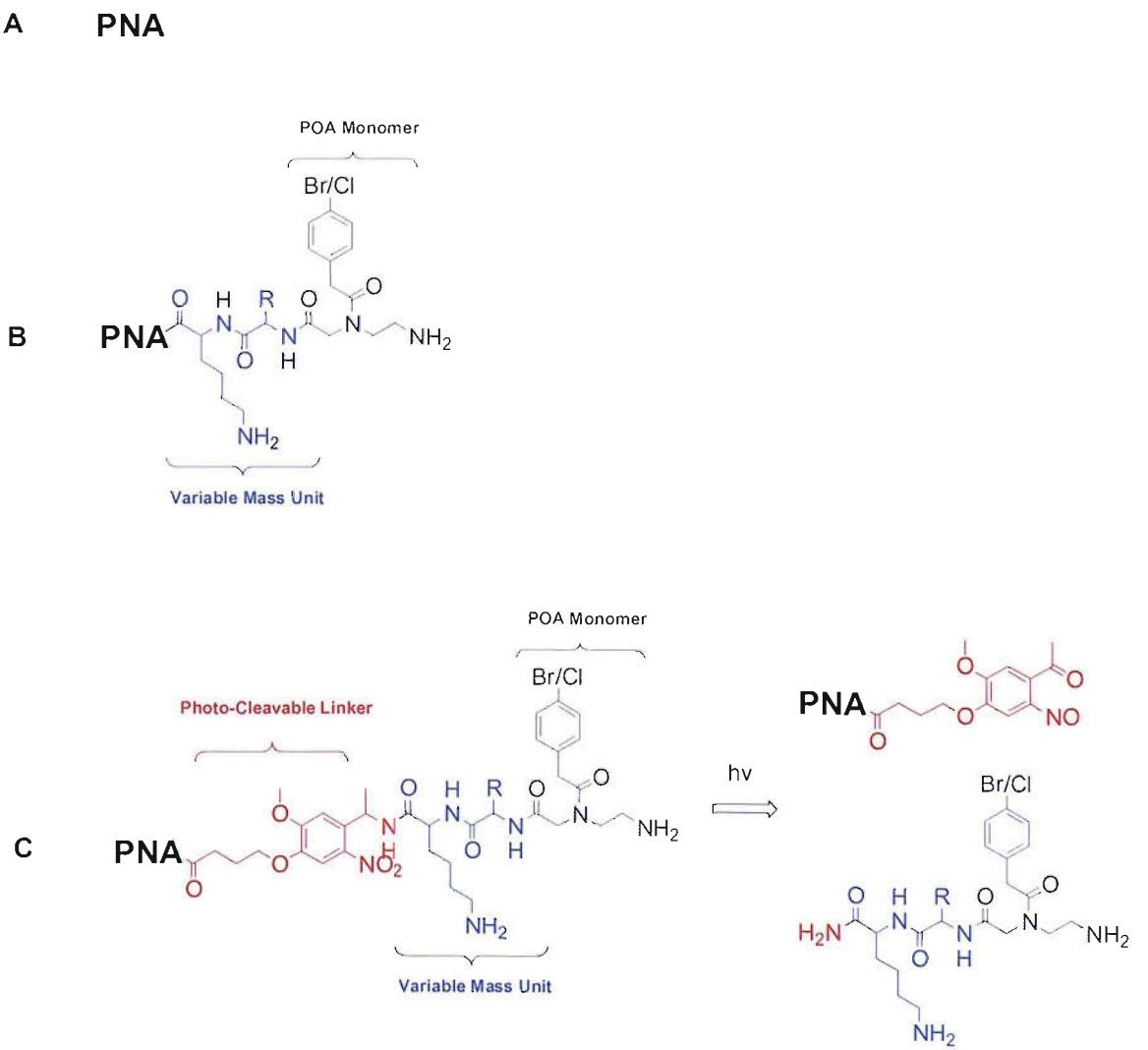


Figure 2.5. The general structure of the PNA probes.

This would allow a direct comparison between the approach using the probes with photo-cleavable mass markers and the existing procedures using unmodified PNA probes[113, 114] and PNA probes modified with mass markers.[115]

As PNA oligomer synthesis is based on the established solid phase peptide chemistry, peptide-like manipulations such as the coupling of compounds containing acid moieties to the amino-terminus of the PNA oligomer are readily performed. This enables the facile synthesis of PNA–peptide conjugates where the PNA section is complementary to the target DNA, whilst the peptide portion serves as a mass marker. This approach is similar to that described by Griffin and co-workers.[115] The smallest possible mass difference between two unmodified PNA probes of equal length is 9 Da, which involves the substitution of an **A** with a **T**. It has been shown that such small mass differences are difficult to resolve on low specification instruments. The function of the mass marker is to enhance the mass difference between probes of similar mass. This would permit unambiguous assignment of the variant base in addition to alleviating the need for extensive sample clean up and a high specification mass spectrometer.

Unique low molecular mass peptidic markers can also be covalently attached to the amino-terminus of the PNA probe *via* a photo-cleavable linker. The marker released is then directly related to the sequence of the PNA probe, which in turn identifies the sequence of the DNA target. Working in this mass region (500-1000 Da) permits the use of monoisotopic resolution without the need for a high-resolution mass spectrometer and in addition permits a greater degree of multiplexing.

To enable facile synthesis of PNA–peptide conjugates all monomers used contained a 9*H*-9-fluorenylmethoxycarbonyl (Fmoc) protected amino functionality and a carboxylic acid moiety to enable all manipulations to be performed on the solid phase.

The mass markers were designed to allow fine-tuning of the molecular weight to provide the flexibility to produce the hundreds of markers needed for multiplex reactions. The mass markers were composed of two sections; a variable mass unit and an isotopically labelled peptide organic acid monomer (POA) in addition to the photo-cleavable linker (**Figure 2.5**).

As a consequence of the un-charged backbone the aqueous solubility of PNA is relatively low compared with DNA. The solubility is dependent on the purine:pyrimidine ratio in addition to the oligomer length. Incorporation of a terminal lysine,[116, 135] an *O*-spacer unit[135] or a solubility enhancer, in which the nucleobase has been replaced with a positively charged hydrophilic moiety,[138] has been shown to improve the solubility. For this reason the variable mass components were composed of lysine to aid solubility and add another low molecular mass amino acid. Amino acids are commercially available and coupling different combinations of amino acids allows the adjustment of the mass of the marker to avoid overlap between markers. Primarily, amino acids such as alanine and phenylalanine, without reactive side chains were chosen to avoid complications during oligomer synthesis.

Finally, POA monomers that are isotopically labeled with either a bromine or chlorine were used. The synthesis of the individual monomers is discussed in detail in **Chapter 4**. The POA monomers were used in conjunction with the variable mass components to give each PNA probe a marker with a unique molecular mass and to maintain the molecular mass of the cleaved marker within the 500 – 1000 Da mass range. The isotope enables the use of a proprietary software package, *e.g.* cluster analysis (MassLynx 4.0), which will automatically identify designated patterns, *e.g.* naturally occurring isotopes or synthetically designed isotope patterns. In addition, an isotope pattern enables facile identification of the mass marker when present in the mass spectrum at low levels and the ions from the mass marker can readily be distinguished from the matrix ions.

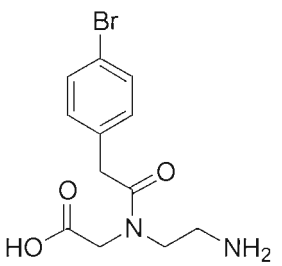
The use of *O*-nitrobenzyl photo-cleavable linkers for the detection of primer extension products by mass spectrometry[139-141] and the cleavage of peptides after solid phase synthesis has been well documented.[142, 143] 4-[4-(1-Aminoethyl)-2-methoxy-5-nitrophenoxy]butanoic acid **(3)** was chosen as the photo-labile linker for incorporation into the PNA probes. The linker is readily cleaved by irradiation at 365 nm[144] which is close to the wavelength of the nitrogen laser in the MALDI source (337 nm). This enables cleavage of the linker and subsequent release of the marker directly from the MALDI target by the nitrogen laser. This alleviates the need for the photo-cleavage of the marker prior to mass spectral analysis, streamlining each analysis and making it amenable to high throughput applications. The alkoxyl group together with the α-methyl on the benzylic carbon have been shown to enhance the rate of cleavage and minimise undesirable side reactions.[144] The linker readily under goes photolysis resulting in the generation of a ketone and an amide (**Figure 2.5**). In addition, the linker has the necessary amino and carboxylic acid functionalities for incorporation into the PNA probe and is commercially available as the Fmoc protected amine.

2.2.4. Synthesis of PNA Probes

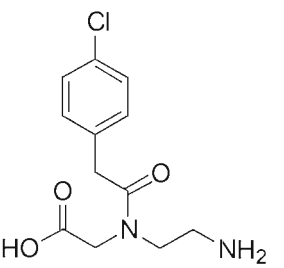
As discussed in the **Chapter 4**, the methods adopted for the assembly of PNA are very similar to those used in the preparation of peptides, which are performed on the solid phase. This enabled the synthesis of three different types of PNA probe; the unmodified PNA probe (PNA), the PNA probe modified with mass marker (PNA-M) and the PNA probe modified with photo-cleavable mass marker (PNA-L-M) from one initial PNA scaffold. For each type of PNA probe two allele specific probes 14 bases in length were prepared to enable the detection of a **G** to **T** mutation on the W1282X locus; one complementary to the wild type allele (PNA(14**C**)) and one complementary to the single base substituted or mutant allele (PNA(14**A**)) (**Table 2.1**).

PNA I.D.	Sequence	Calculated Average Mass (Da) $(M+H)^+$	Calculated Monoisotopic Mass (Da) $(M+H)^+$
PNA(14**C**)	CTTTCCT**C**CACTGT	3690.5740	3688.4595
PNA(14**A**)	CTTTCCT**A**CACTGT	3714.5991	3712.4707
PNA-M(14**C**)	Br-Ala-Lys-CTTTCCT**C**CACTGT	4186.9785	4183.6076
PNA-M(14**A**)	Cl-Phe-Lys-CTTTCCT**A**CACTGT	4242.6501	4239.7007
PNA-L-M(14**C**)	Br-Ala-Lys-hv-CTTTCCT**C**CACTGT	4467.2590	4463.7137
PNA-L-M(14**A**)	Cl-Phe-Lys-hv-CTTTCCT**A**CACTGT	4522.9306	4519.8066

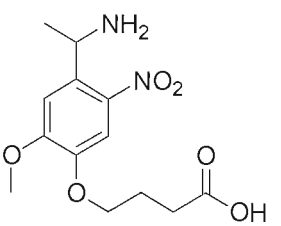
Table 2.1. PNA probes synthesised for the pilot study. Br = 2-(2-aminoethyl)[2-(4-bromophenyl)acetyl]aminoacetic acid (POA monomer (1)), Cl = 2-(2-aminoethyl)[2-(4-chlorophenyl)acetyl]aminoacetic acid (POA monomer (2)) and hv = 4-[4-(1-aminoethyl)-2-methoxy-5-nitrophenoxy]butanoic acid (photo-cleavable linker (3)) (Figure 2.6).



1



2



3

Figure 2.6. The structures of the POA monomers and photo-cleavable linker.

2.2.5. UV-Melting

Duplex stability and determination of the T_m is important in probe design. In their respective PNA hybridisation assays Griffin and Jiang-Baucom used a buffer composed of Tris and NaCl at pH 7 to achieve the best discrimination between fully complementary and single base mismatched PNA/DNA duplexes.[114, 115] The PNA probes used in those assays were of different lengths so the exact buffer conditions could not be directly transferred to the assay described here.

In order to identify the optimum concentration of NaCl in the hybridisation buffer required to obtain a functional T_m, the hybridisation behavior of the unmodified wild type PNA probe (PNA(14**C**)) was assessed by UV melting with complementary and single base mismatched synthetic DNA targets in a 10 mM Tris buffer with varying concentrations of NaCl. The T_m (**Table 2.2**) values were determined from the maximum of the first order derivative (average of the three consecutive heat and cool cycles). Analysis of the data showed excellent discrimination between fully matched and single base mismatch PNA/DNA duplexes at all concentration of NaCl. Although there was little difference in the T_m values observed using a buffer with 0.5 M and 1 M NaCl, a buffer composed of 10 mM Tris, 1 M NaCl pH 7 was chosen as the hybridisation buffer in the assay.

PNA I.D.	DNA Target Sequence	Base Pair	[NaCl]	T_m (°C)	ΔT_m (°C)
PNA(14**C**)	ACAGTG**G**AGGAAAG	**C**--**G**	0.0	ND	-
PNA(14**C**)	ACAGTG**T**AGGAAAG	**C**--**T**	0.0	52.3	-
PNA(14**C**)	ACAGTG**G**AGGAAAG	**C**--**G**	0.1	62.2	-
PNA(14**C**)	ACAGTG**T**AGGAAAG	**C**--**T**	0.1	44. 6	17.6
PNA(14**C**)	ACAGTG**G**AGGAAAG	**C**--**G**	0.5	54.8	-
PNA(14**C**)	ACAGTG**T**AGGAAAG	**C**--**T**	0.5	36.8	18.0
PNA(14**C**)	ACAGTG**G**AGGAAAG	**C**--**G**	1.0	54.0	-
PNA(14**C**)	ACAGTG**T**AGGAAAG	**C**--**T**	1.0	35.5	18.5

Table 2.2. T_m values for the unmodified PNA probe PNA(14C) with synthetic DNA targets in 10 mM Tris with different concentration of NaCl at pH 7. ΔT_m is the difference in T_m between the fully matched and single base mismatch PNA/DNA duplexes.

In order to investigate the ability of the modified and unmodified PNA probes to discriminate between fully complementary and single base mismatched DNA targets, thermal denaturation experiments were carried out with the PNA probes and synthetic oligodeoxynucleotides (**Figure 2.7**). The probes showed excellent discrimination between the perfectly matched and single base mismatched alleles with T_m values differing by up to 20.6 °C (**Table 2.3**) between the wild-type PNA probe (PNA(14**C**)) and mutant DNA target (**T**). In addition, little variation was observed in the stability of the modified and unmodified PNA/DNA duplexes. This demonstrates that the linker and mass markers have little effect on the stability of the PNA/DNA duplex and as a consequence the same assay conditions can be used with the modified and unmodified PNA probes.

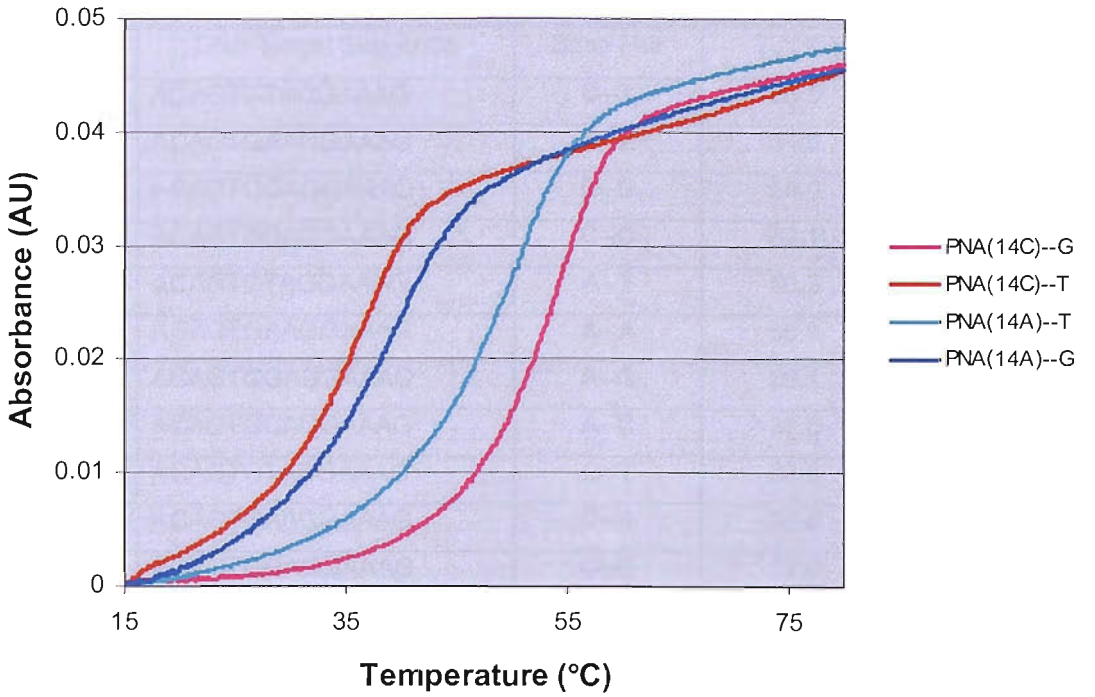


Figure 2.7. UV-Melting profiles of the unmodified PNA probes PNA(14C) and PNA(14A) with synthetic DNA targets in 10 mM Tris, 1 M NaCl, pH 7.

PNA I.D.	DNA Target Sequence	Base Pair	T_m (°C)	ΔT_m (°C)
PNA(14**C**)	ACAGTG**T**AGGAAAG	**C**--T	35.5	18.5
PNA(14**C**)	ACAGTG**A**AGGAAAG	**C**--A	34.6	19.4
PNA(14**C**)	ACAGTG**G**AGGAAAG	**C**--G	54.0	-
PNA(14**C**)	ACAGTG**C**AGGAAAG	**C**--C	34.1	19.9
PNA(14**A**)	ACAGTG**T**AGGAAAG	**A**--T	50.3	-
PNA(14**A**)	ACAGTG**A**AGGAAAG	**A**--A	36.0	14.3
PNA(14**A**)	ACAGTG**G**AGGAAAG	**A**--G	39.1	11.2
PNA(14**A**)	ACAGTG**C**AGGAAAG	**A**--C	37.6	12.7
PNA-M(14**C**)	ACAGTG**T**AGGAAAG	**C**--T	33.8	18.2
PNA-M(14**C**)	ACAGTG**A**AGGAAAG	**C**--A	31.8	20.2
PNA-M(14**C**)	ACAGTG**G**AGGAAAG	**C**--G	52.0	-
PNA-M(14**C**)	ACAG**T**G**C**AGGAAAG	**C**--C	36.5	15.5
PNA-M(14**A**)	ACAG**T**G**T**AGGAAAG	**A**--T	50.4	-
PNA-M(14**A**)	ACAG**T**G**A**AGGAAAG	**A**--A	32.2	18.2
PNA-M(14**A**)	ACAG**T**G**G**AGGAAAG	**A**--G	37.6	12.8
PNA-M(14**A**)	ACAG**T**G**C**AGGAAAG	**A**--C	37.7	12.7
PNA-L-M(14**C**)	ACAGTG**T**AGGAAAG	**C**--T	34.4	18.4
PNA-L-M(14**C**)	ACAGTG**A**AGGAAAG	**C**--A	32.3	20.5
PNA-L-M(14**C**)	ACAGTG**G**AGGAAAG	**C**--G	52.8	-
PNA-L-M(14**C**)	ACAGTG**C**AGGAAAG	**C**--C	35.7	17.1
PNA-L-M(14**A**)	ACAGTG**T**AGGAAAG	**A**--T	51.7	-
PNA-L-M(14**A**)	ACAGTG**A**AGGAAAG	**A**--A	39.0	12.7
PNA-L-M(14**A**)	ACAGTG**G**AGGAAAG	**A**--G	34.1	17.6
PNA-L-M(14**A**)	ACAG**T**G**C**AGGAAAG	**A**--C	39.5	12.2

Table 2.3. T_m values for the three different types of PNA probe with synthetic DNA targets in 10 mM Tris, 1 M NaCl, pH 7. ΔT_m is the difference in T_m between the fully matched and single base mismatch PNA/DNA duplexes.

2.2.6. Assay conditions

The proposed PNA hybridisation assay comprises a series of molecular recognition events, culminating in the release and detection of the mass marker by MS. As a general strategy it was considered simpler to immobilise the double stranded biotinylated PCR product and chemically denature the duplex leaving a single stranded DNA target as described in **Figure 2.8** (**Figure 1.22**) as opposed to thermally denaturing the PCR product in solution and immobilising the PNA/DNA duplex.

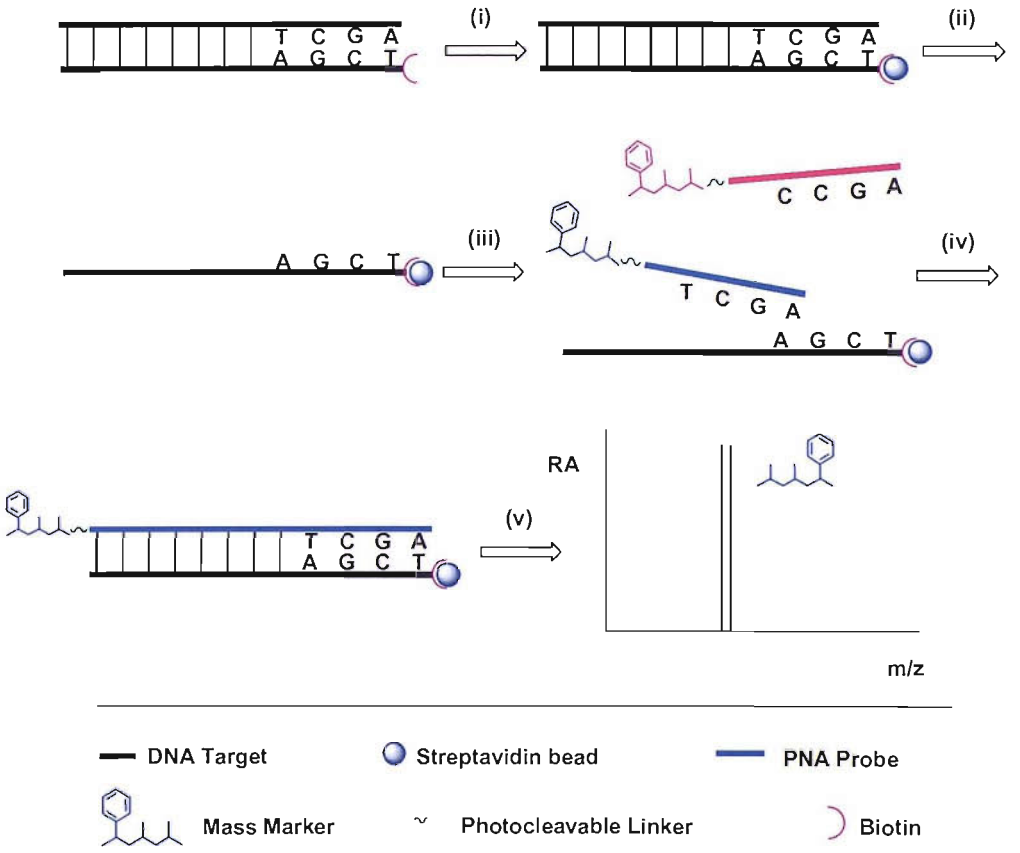


Figure 2.8. PNA hybridisation assay. (i) A double stranded biotinylated PCR product is immobilised on a streptavidin bead. (ii) The duplex is denatured and the unbound strand is washed away. (iii) Modified PNA probes complementary to the two possibe alleles are added. (iv) Uncomplementry probes are removed by heating and washing. (v) The PNA/DNA/bead complex is analysed by MALDI-TOF-MS. The genotype of the sample is determined by the quasi-molecular ion of the mass marker cleaved from the PNA probe.

The beads used in the assay are hydrophobic, 2.8 μm in diameter, paramagnetic and coated in a monolayer of covalently conjugated streptavidin. The interaction between the bead and the target is based on the high affinity of the streptavidin ligand on the surface of the Dynabeads for the biotin on the DNA target. The beads are superparamagnetic, *i.e.*, they exhibit magnetic properties only when placed within a magnetic field and show no residual magnetism when removed from the field. Consequently the target/bead complex can be removed from the suspension using a magnet. Bead/target complexes are drawn to the side of the tube facing the magnet, allowing the supernatant to be removed with a pipette. The inherent benefits of magnetic handling allow for easy washing, separation and concentration of the target without the need for centrifugation or columns.

A basic assay protocol was initially established with a single stranded biotinylated synthetic oligonucleotide, which was designed to mimic the wild type PCR product. Using the single stranded target alleviates the need for the chemical denaturation step thus simplifying the assay and enables a known quantity of DNA target to be immobilised on the beads. Initially, 30 μg of beads were washed with aliquots of binding buffer (10 mM Tris, 1 M NaCl, pH 7) as described by Griffin *et al.*[111] The beads were re-suspended in the binding buffer and 10 pmol of the single stranded synthetic DNA target were added to the solution. The solution was left for 20 minutes at room temperature and the beads were subsequently washed with aliquots of the binding buffer. Again the beads were re-suspended in the same buffer and 30 pmol of wild type PNA probe (PNA(14**C**)) and 30 pmol of the mutant probe (PNA-M(14**A**)) were added. The beads were incubated at 80 °C, cooled to room temperature and washed with 10 mM Tris, 0.1 % SDS, pH 7 (wash buffer) and water. The beads were re-suspended in the minimum quantity of matrix (α-cyano-4-hydroxycinnamic acid in acetonitrile:water 50:50 (0.1 % TFA)) and spotted directly on the MALDI target with a few Dowex 50WX8-200 ion exchange beads to remove any residual sodium ions.[57]

It was reported by Griffin and co-workers that the PNA probe was denatured from the immobilised DNA during the application of the matrix allowing the analysis of the PNA probe by MALDI-TOF-MS.[115]

The wild type PNA probe (PNA(14**C**)) was readily detected with minor traces of the mutant probe (PNA-M(14**A**)) when analysed by positive ion linear MALDI-TOF-MS. The results showed a second incubation step is essential to fully discriminate between the fully complimentary probe and the probe with a single base mismatch. In addition, blank samples with no DNA target showed the non-specific binding of the PNA probes directly to the coating of the bead. 0.1 % BSA was added to the binding buffer during hybridisation to prevent this interaction.

It was necessary to determine the temperature at which there is discrimination between the fully complimentary probe and the probe with a single base mismatch without compromising sensitivity. In order to identify the optimum incubation temperature the assay was performed with a second incubation step after the initial heating at 80 °C incubating the beads to 35, 40, 45, 50, 55 and 60 °C for 20 minutes in the binding buffer. Results showed that incubation at 45 °C gave the optimal balance between the hybridisation selectivity and sensitivity.

2.2.7. Limit of Determination of the Assay

Having established a reliable experimental protocol it was necessary to determine the limitations of the assay with respect to sensitivity. The limit of determination for the assay was established for each type of probe with the single stranded synthetic wild type DNA target. The binding capacity of the beads is fragment length dependent due to steric factors. As a result a larger quantity of the short single stranded DNA target would be immobilised than a double stranded PCR product.

Whilst using a short synthetic target is not a truly accurate assessment of the sensitivity of the assay it enables the immobilisation of a known quantity of DNA target to the beads and allows a direct comparison of the sensitivity of the three different types of PNA probe. The limit of determination was established as follows; a known concentration of DNA target was added to 30 μg of streptavidin beads. After washing, 30 pmol of the wild type and mutant PNA was added and the mixture was left at room temperature for 15 minutes. The mixture was heated to 80 °C for 1 hour, cooled to room temperature, washed and incubated at 45 °C for 20 minutes. The beads were washed and loaded directly on the MALDI platen for analysis in positive ion linear mode (PNA and PNA-M) and positive ion reflectron mode (PNA-L-M). The limit of determination of the assay was defined as the lowest concentration of DNA target at which the wild type PNA probe could be observed in the mass spectrum (**Figure 2.9**).

The limits of determination for the assay using the unmodified PNA probes (PNA(14**C**) **and (PNA(14A**)) (**Figure 2.9**) and the PNA probes modified with mass markers (PNA-M(14**C**) and (PNA-M(14**A**)) were both 100 fmol of single stranded DNA target. This further supports the UV melting data and demonstrates that the mass markers do not affect the binding and specificity of the PNA probes in the assay. The mass marker, cleaved by the nitrogen laser from the immobilised PNA probe, could not be detected even when 100 pmol of DNA was added to the beads but the intact probe was detected. It is worth noting that the binding capacity of 30 mg of beads is less than 7 pmol of single stranded DNA.

The mass spectrum in **Figure 2.10** was acquired in reflectron mode with a high laser power in order to maximise the cleavage of the mass marker. Under these conditions ions corresponding to the intact probe and degradation products were observed with 10 pmol of DNA target.

This data implies that under the instrumental conditions used, the efficiency of the cleavage of the linker is insufficient to readily detect the mass marker cleaved from the PNA probe.

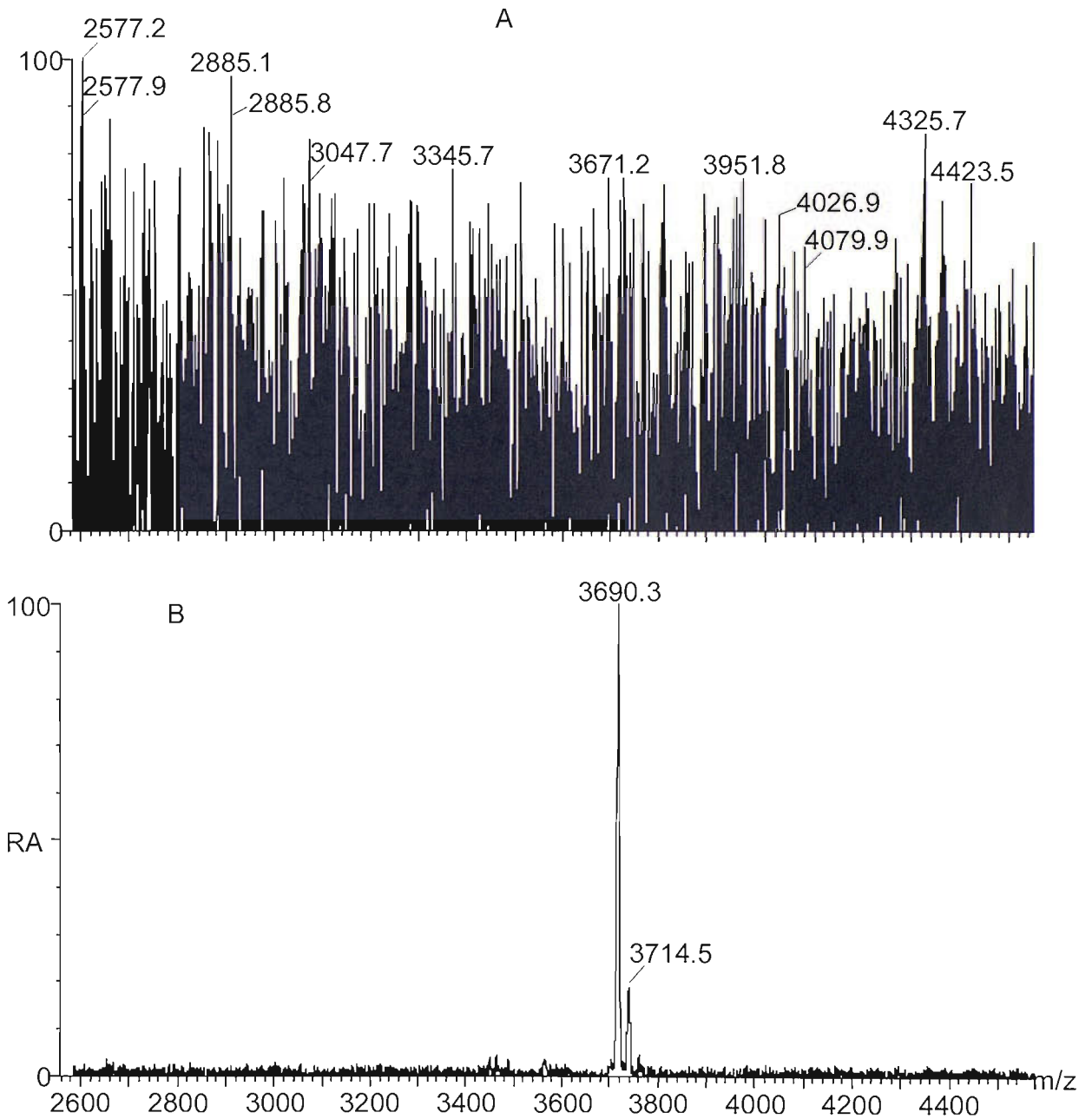


Figure 2.9. Linear MALDI-TOF mass spectrum acquired in linear mode of the unmodified PNA probe (PNA(14C)) released from the immobilised single stranded DNA target at 10 fmol (A) and 100 fmol (B). The limit of determination was 1 pmol. Expected mass $(M+H)^+$ 3690.6 Da.

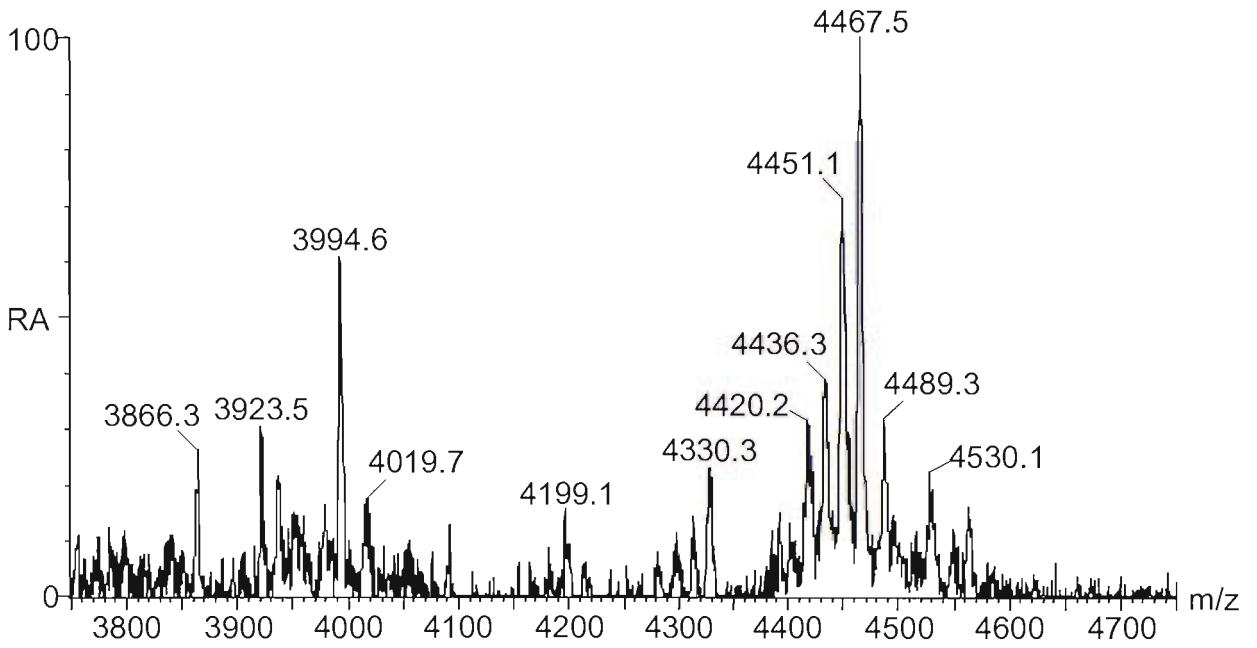
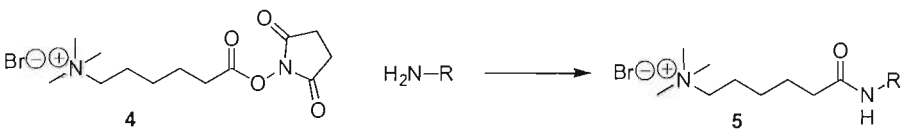


Figure 2.10. Reflectron MALDI-TOF mass spectrum of the PNA probe modified with photo-cleavable mass marker (PNA-L-M(14C)) released from the immobilised single stranded DNA target. Expected mass $(M+H)^+$ 4467.2 Da. Additional ions are observed due to the cleavage of the photo-cleavable linker.

2.2.7.1. Derivatisation of the Mass Marker

One of the advantages of using peptide based mass markers is that manipulations analogous to those used on peptide and proteins can be readily performed. The introduction of a charged moiety on peptides for increased sensitivity[145] and to promote and control fragmentation[146, 147] has been extensively studied. The charge can be introduced on the N-terminus, C-terminus or side chain. Bartlet-Jones *et al.* described the synthesis of a quaternary ammonium activated ester (C_5Q) **(4)**, which was readily attached to the amino terminus of peptides (**Scheme 2.1**).[145] The derivatisation was performed on 50 fmol of peptide directly on the MALDI target in a volatile buffer.

This derivatisation approach was later used by Spengler[146] to control the fragmentation of peptides in PSD MALDI-TOF-MS and Gut *et al.* to enhance the efficiency of detection of modified DNA by MALDI-TOF-MS.[103, 106]



Scheme 2.1 Derivatisation of the amino terminus of a peptide with C_5Q.

Derivatisation with the charge tag C_5Q **(4)** is an ideal means of introducing a preformed positive charge onto the mass marker as it s readily synthesised (the synthesis of C_5Q is discussed in **Chapter 4**), it does not require toxic reagents such as CH_3I[148] and can be incorporated in a volatile buffer on the MALDI target. In an attempt to assess the feasibility of using C_5Q to derivatise the mass marker covalently linked to the immobilised PCR product, C_5Q was added to PNA-L-M(14**C**) on the MALDI target. The buffer was removed *in vacuo* and the MALDI matrix added. Analysis by mass spectrometry showed extra ions containing a bromine isotope pattern (m/z 764.5, 809.5 and 823.5) in addition to the expected product (M^+ expected m/z 668.3 Da) (**Figure 2.11**). In their work on the derivatisation of peptides Spengler *et al.* discussed the feasibility of a doubly C_5Q derivatised species but observed solely the singly charged species $(M+2C_5Q-H)^+$. They concluded one of the two fixed charges is neutralised internally by the formation of a zwitterion.[146] In this case it is probable that in addition to derivatising the amino terminus the amino functionality on the lysine side chain has also been derivatised with C_5Q (**Figure 2.12**). It is of particular interest that only the singly charged species is observed in the mass spectrum. It is probable that the ions observed at m/z 809.5 and 709.5 are fragment ions and are the result of the loss of CH_2 and $N(CH_3)_3$ respectively.

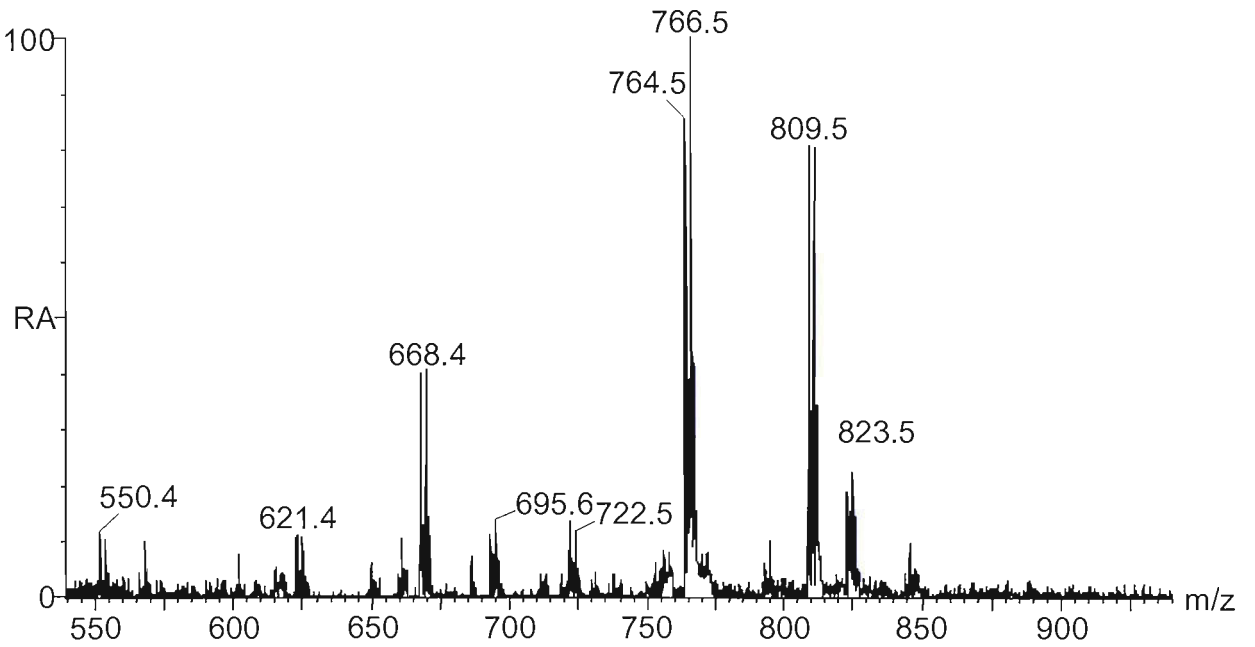


Figure 2.11. Reflectron MALDI-TOF mass spectrum of the derivatised mass marker photochemically released from the wild type PNA probe modified with photo-cleavable mass marker (PNA-L-M(14C)). Expected mass $(M)^+$ 668.3 Da. Additional ions were observed due to uncontrolled derivatisation.

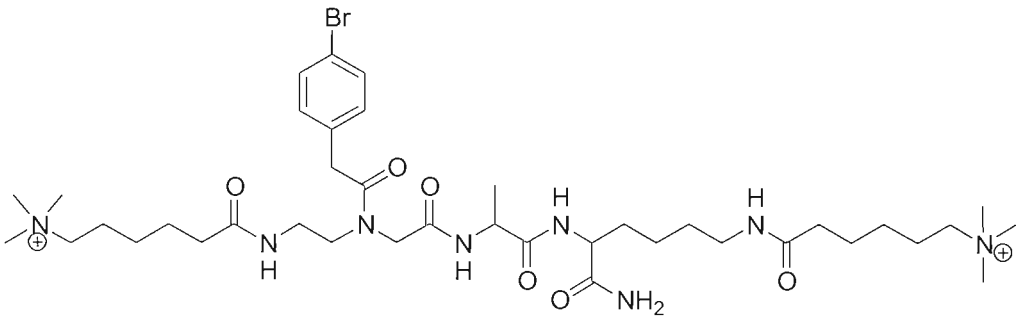


Figure 2.12. Doubly derivatised mass marker cleaved from PNA probe PNA-L-M(14C).

In an attempt to identify the conditions under which the positive charge is selectively introduced onto one amino functionality 250 pmol PNA-M-L(14C) was derivatised with decreasing equivalents of C_5Q. Analysis by MALDI-TOF-MS showed a ratio of 5:1 PNA:C_5Q was optimum to derivatise only one functionality (**Figure 2.13**). No further limit of determination experiments to determine the effect of the positive charge were carried out as the limit of determination would be established with the next batch of PNA probes synthesised.

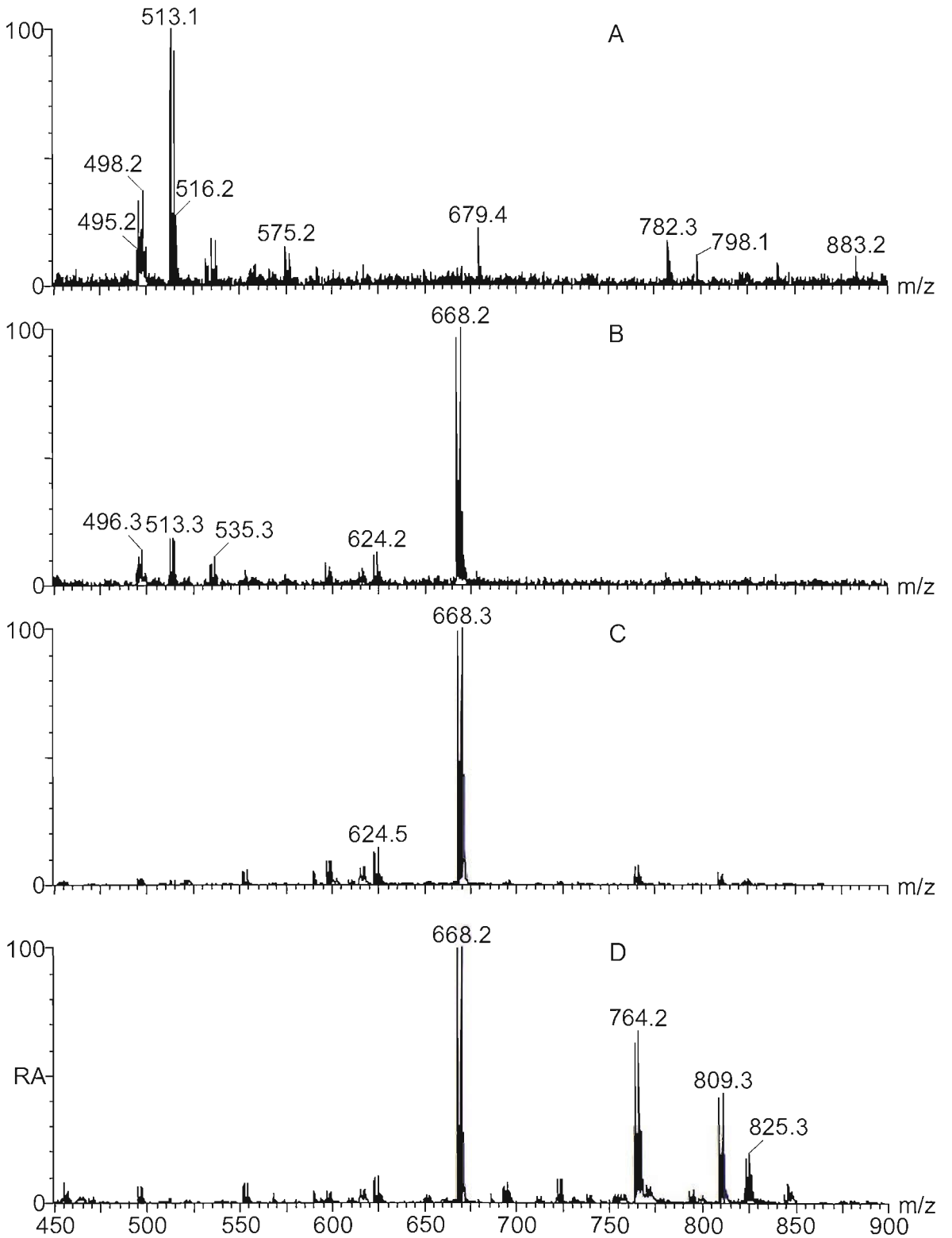


Figure 2.13. Derivatisation of 250 pmol of PNA-M-L(14C) with decreasing equivalents of C_5Q. Ratio of PNA:C_5Q; (A) 50:1, (B) 5:1, (C) 1:1 and (D) 1:2.

2.2.8. PCR of Human DNA

The W1282X locus of the wild type human genomic DNA was amplified by PCR. Sequence data for the ABCC7 locus was obtained from GenBank. The W1282X primers were designed to give amplicons of 100 bases with the forward primer being modified with a 5'-biotin.[45, 149] PCR conditions were adapted from those described by Thewell and co-workers.[45, 149] All PCR reactions were performed in parallel with no template controls in which no DNA template was added to the reaction. In addition, aliquots of the samples were analysed by gel electrophoresis with respect to a 25 base pair DNA stepladder prior to purification to confirm the success of the PCR reaction.

The free biotinylated PCR primers bind to the Dynabeads much more rapidly than longer PCR fragments, so it is necessary to ensure that all unextended primers are removed prior to the immobilisation step. To prevent an excess of free biotinylated primers, it is possible to perform PCR with a low concentration of the biotinylated primer, or to remove the excess primers by purification after PCR. In order to get the maximum yield of PCR amplicon an excess of primers was used in the PCR reaction and the products were subsequently purified on silica spin columns. Purification has the added advantage of removing the enzyme, detergents and other components used in the PCR reaction as well as concentrating the sample. **Figure 2.14** shows a photograph of a gel stained with ethidium bromide of the unpurified PCR sample, unpurified no template control, the purified sample and purified no template control with respect to a 25 base pair DNA stepladder. The gel image demonstrates that no PCR occurs in the no template controls and that minimal sample is lost during purification.

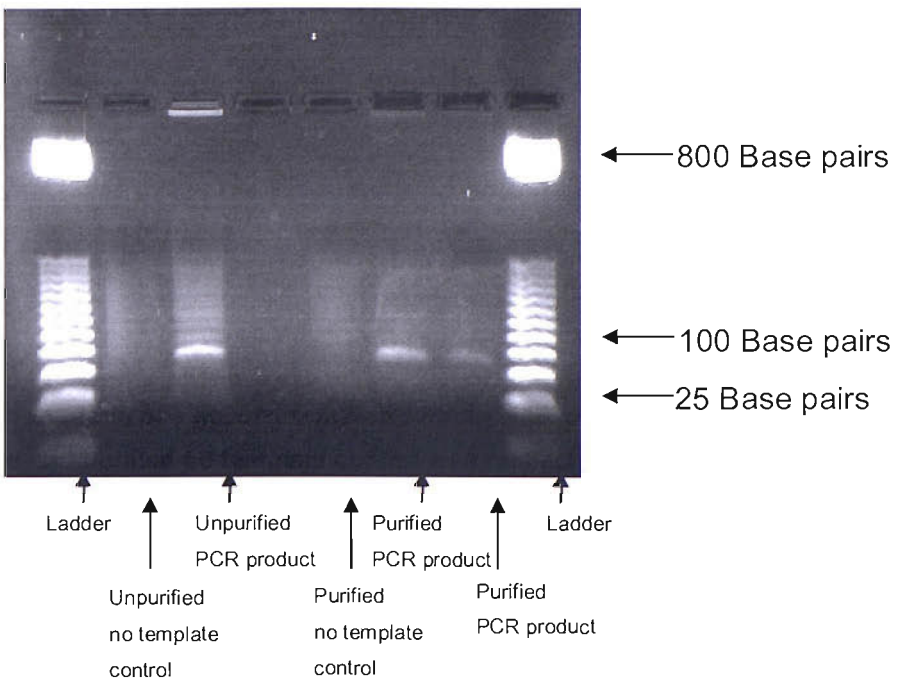


Figure 2.14. Photograph of a gel of unpurified sample and no template control and the purified sample and purified no template control with respect to a 25 base pair DNA stepladder.

2.2.9. Assay with PCR Amplified Human DNA

The three different types of probe were used in the PNA hybridisation assay to simulate a **G** to **T** mutation in the W1282X locus. The assay was carried out as described previously. However, PCR amplicons generated from wild type human DNA were used in place of the single stranded DNA targets. In addition, an extra step to chemically denature the immobilised PCR amplicon was introduced.

Initially the assay was performed in triplicate with the unmodified PNA probes in parallel with a blank sample (*i.e.* no PCR product). Following immobilisation on the beads, the double stranded PCR amplicon was denatured with sodium hydroxide and the unbound strand washed away. An excess of the two PNA probes representing the wild type (PNA(14**C**)) and mutant (PNA(14**A**)) alleles were added. The solution was heated and cooled to allow the fully complementary probes to hybridise to the DNA target. Following washing, the solution was incubated one further time to remove any remaining uncomplementary probes. The beads were washed to remove any salts and impurities and the PNA/DNA/bead complex was loaded directly on the MALDI target with the matrix. As can be seen in **Figure 2.15**, only the wild type (PNA(14**C**)) PNA probe was detected in linear mode by positive ion MALDI-TOF-MS.

The assay was repeated with the PNA probes modified with mass markers. Exclusively the wild type (PNA-M(14**C**)) PNA probe was detected in linear mode by positive ion MALDI-TOF-MS (**Figure 2.16**). The mass markers increase the mass difference between the wild type and mutant probes to 56.1 Da opposed to 24.0 Da when using the unmodified probes. This alleviates any ambiguity when salt adducts are observed in the mass spectrum.

Finally the assay was evaluated with the PNA probes modified with photo-cleavable mass markers (PNA-L-M(14**C**)). It was difficult to control the derivatisation of the mass marker with C_5Q on the MALDI target due to different amounts of PNA immobilised on the single stranded PCR amplicon. However, **Figure 2.17** shows derivatisation of the mass marker with C_5Q prior to analysis by mass spectrometry allowed facile detection of the mass marker released from the PNA probe during laser ablation.

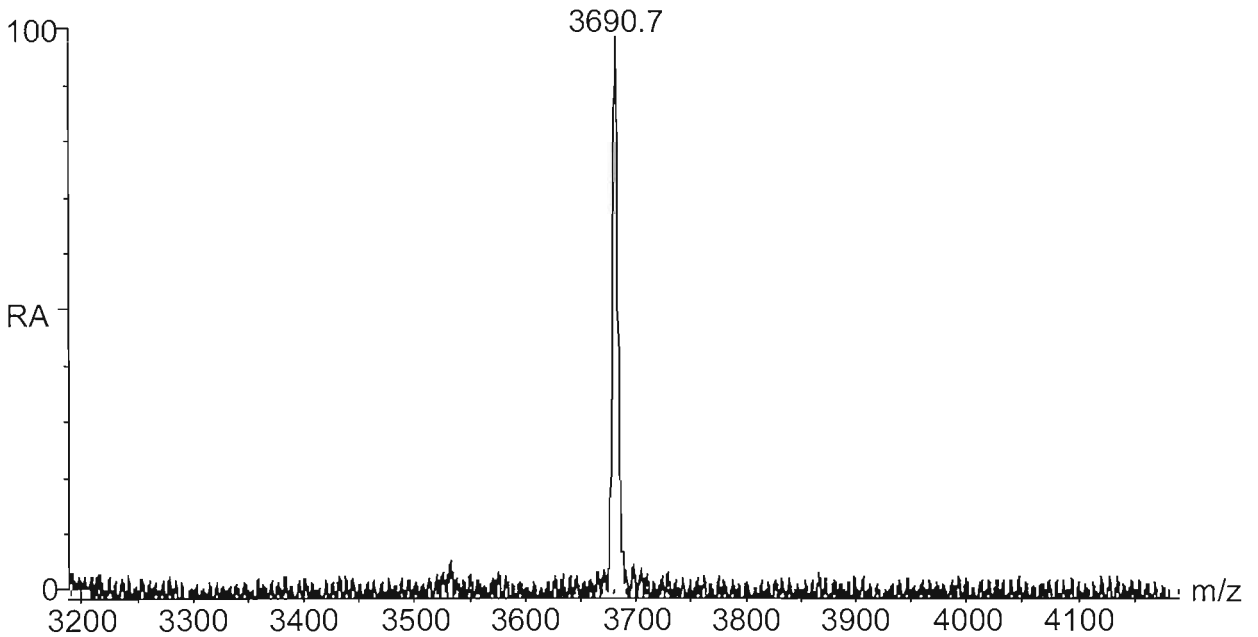


Figure 2.15. Linear MALDI-TOF mass spectrum of the unmodified PNA probe (PNA(14C)) released from the immobilised PCR product. Expected mass $(M+H)^+$ 3690.6 Da.

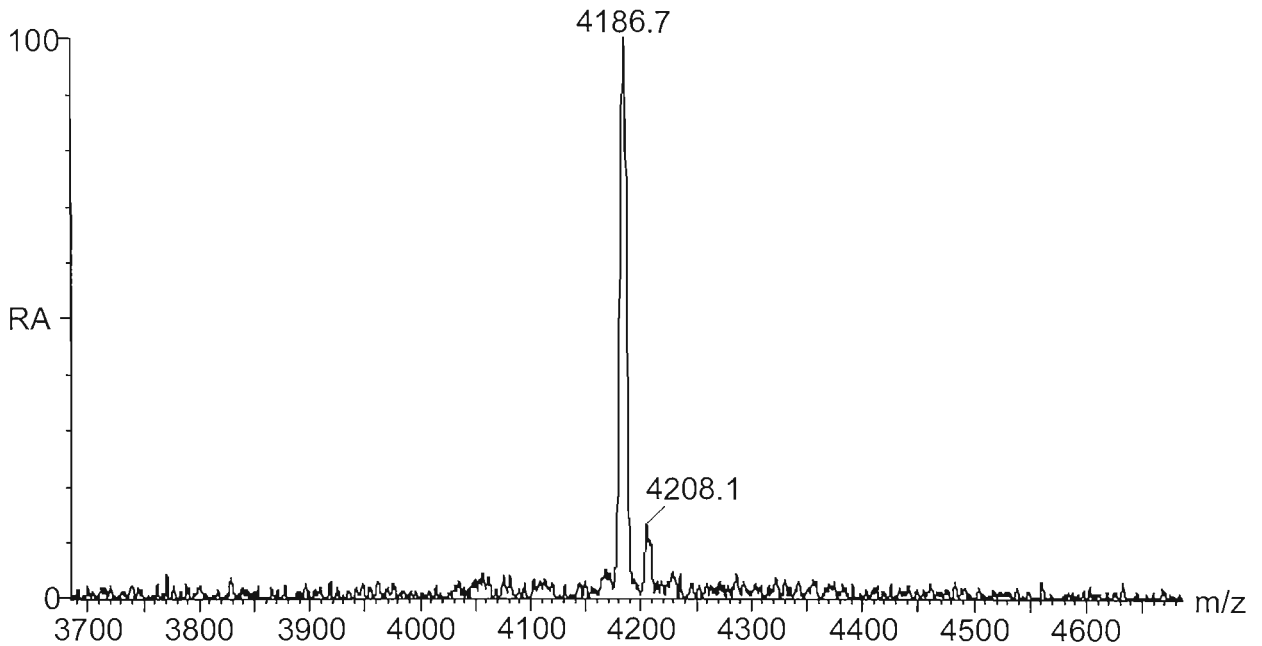


Figure 2.16. Linear MALDI-TOF mass spectrum of the PNA probe modified with mass marker (PNA-M(14C)) released from the immobilised PCR product. Expected mass $(M+H)^+$ 4187.0 Da.

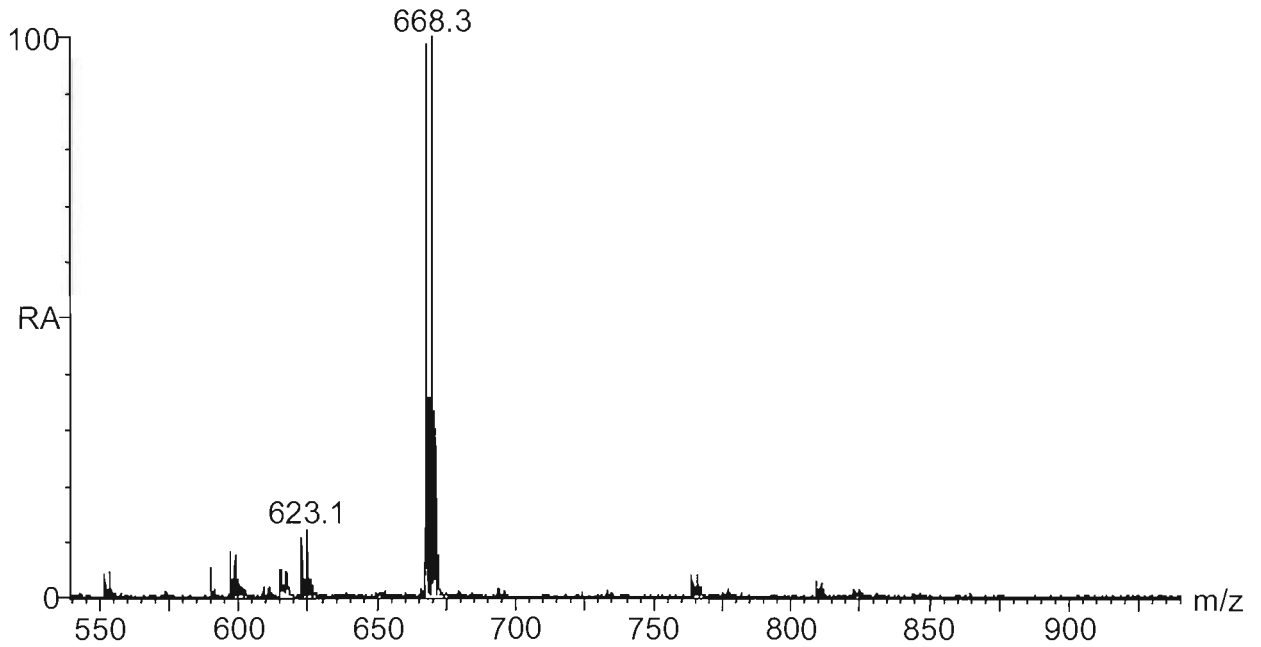


Figure 2.17. Reflectron MALDI-TOF mass spectrum of the mass marker photochemically released from the wild type PNA probe modified with photo-cleavable mass marker (PNA-L-M(14C)). Expected mass $(M)^+$ 668.3 Da.

In order to determine whether the preformed positive charge was stable to the assay conditions and as a result could be incorporated into the PNA probes during solid phase synthesis the charge was introduced onto the amino terminus of the unmodified PNA probes. The PNA was dissolved in 1 M trimethylammonium bicarbonate pH 8.5 at 5 °C and an excess of C_5Q was added. The mixture was left for 2 hours at 5 °C then the samples were dried and purified by RP-HPLC to give one further type of PNA probe, PNA^+ (**Table 2.4**).

The ability of the charged tagged PNA probes to discriminate between fully complementary and single base mismatched DNA targets was determined by UV-melting. Experiments were carried out as described previously with synthetic DNA targets (**Table 2.5**). The probes showed excellent discrimination between the perfectly matched and single base mismatched alleles. In addition, the results demonstrate that the positive charge has little effect on the stability of the PNA/DNA duplex.

PNA I.D.	Sequence	Calculated Average Mass (Da) $(M)^+$	Calculated Monoisotopic Mass (Da) $(M)^+$
PNA$^+$(14**C**)	($(CH_3)_3{}^+N(CH_2)_5CO$)CTTTCCT**C**CACTGT	3845.8142	3843.5905
PNA$^+$(14**A**)	($(CH_3)_3{}^+N(CH_2)_5CO$)CTTTCCT**A**CACTGT	3869.8393	3867.6018

Table 2.4. Charge tagged PNA probes.

PNA I.D.	Oligonucleotide Sequence	Base Pair	T_m (°C)	ΔT_m
PNA$^+$(14**C**)	ACAGTG**T**AGGAAAG	**C**--**T**	34.7	17.2
PNA$^+$(14**C**)	ACAGTG**A**AGGAAAG	**C**--**A**	30. 9	21.0
PNA$^+$(14**C**)	ACAGTG**G**AGGAAAG	**C**--**G**	51.9	-
PNA$^+$(14**C**)	ACAGTG**C**AGGAAAG	**C**--**C**	36.7	15.2
PNA$^+$(14**A**)	ACAGTG**T**AGGAAAG	**A**--**T**	49.6	-
PNA$^+$(14**A**)	ACAGTG**A**AGGAAAG	**A**--**A**	38.3	11.3
PNA$^+$(14**A**)	ACAGTG**G**AGGAAAG	**A**--**G**	33.9	15.7
PNA$^+$(14**A**)	ACAGTG**C**AGGAAAG	**A**--**C**	40.0	9.6

Table 2.5. T_m values for charge tagged PNA probes with synthetic DNA targets in 10 mM Tris, 1 M NaCl, pH 7. ΔT_m is the difference in T_m between the fully matched and single base mismatch PNA/DNA duplexes

The hybridisation assay was repeated using PCR amplified DNA and the charge tagged PNA. As can be seen in **Figure 2.18** solely the wild type PNA probe was observed in the MALDI-TOF mass spectrum. This confirms the preformed positive charge was stable to the assay conditions and can be included during the synthesis of the probes.

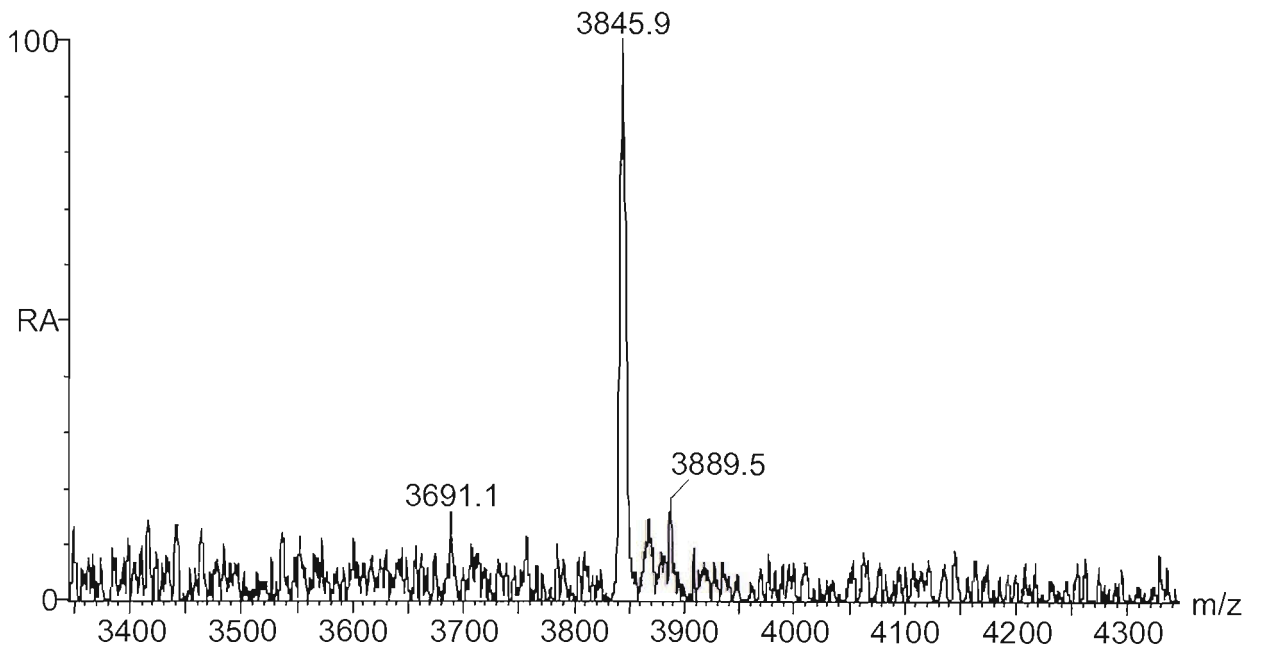


Figure 2.18. Linear MALDI-TOF mass spectrum of the charge tagged wild type PNA probe modified with photo-cleavable mass marker (PNA^{+}(14C)). Expected mass $(M)^{+}$ 3845.8 Da.

2.3. Summary

The basic experimental protocol for a PNA hybridisation assay coupled with detection by MS has been established. Four different types of PNA probe have been synthesised to simulate a **G** to **T** transversion mutation of the W1282X locus: an unmodified PNA probe, a PNA probe modified with a preformed positive charge, a PNA probe modified with mass marker and a PNA probe modified with a photo-cleavable mass marker. The mass markers were composed of inexpensive, commercially available amino acids and an isotopically labelled POA monomer, which could be readily conjugated to the amino terminus of the PNA probe during synthesis on the solid phase. The mass markers enhanced the mass difference between each probe and served as a unique molecular mass identifier in the case of the PNA probe modified with a photo-cleavable mass marker. Each step of the assay and assay conditions were optimised so they could be readily transferred to an automated liquid handling system, rendering it amenable to high throughput applications. In addition, generic instrument parameters were used in order to ensure the assay required minimal optimisation and was compatible with low specification MALDI-TOF instruments.

Thermal denaturation experiments were carried out with single stranded synthetic DNA targets to assess the selectivity of each type of PNA probe. The probes showed excellent discrimination between the fully complementary and single base mismatch DNA targets. The results from the UV-melting demonstrated the modifications to the PNA probes had little or no effect on the stability of the hybrid duplexes.

The sensitivity of the hybridisation assay was evaluated with biotinylated single stranded DNA targets. The results further confirmed the UV-melting results, which suggested that the modification to the PNA probes have little effect on the stability of the PNA/DNA duplex with the limit of determination of the assay being 1 pmol for the unmodified PNA probe, and PNA probe modified with mass marker. Under the generic assay conditions the mass marker cleaved from the PNA probe with the nitrogen laser could not be detected. A protocol, which effectively introduces a preformed positive charge into the mass marker has been established.

Finally, the assay was evaluated with wild type PCR amplified DNA. The unmodified PNA probe and PNA probe modified with mass marker showed excellent selectivity and were readily detected with good mass accuracy by MALDI-TOF-MS. It was necessary to derivatise the mass marker with a quaternary ammonium moiety to detect the mass marker cleaved from the PNA probe during laser ablation.

In conclusion, the feasibility of using four different types of PNA probe in a hybridisation assay for the detection of a **G** to **T** mutation of the W1282X locus was investigated. The probes have shown excellent selectivity for the complementary region of amplified DNA. Modification of the mass marker to introduce a preformed positive charge was essential for the successful detection of the mass marker cleaved from the PNA probe.

3. PNA Hybridisation Assay

3.1. Synthesis of PNA Probes

The following four types of PNA probe were prepared for use in the hybridisation assay to detect the cystic fibrosis W1282X mutation (**G** to **A**) (**Table 3.1**, **Figure 3.1**)

- Unmodified PNA probe (**A**).
- Charge tagged PNA probe (**B**).
- PNA probe modified with a charge tagged mass marker (**C**).
- PNA probe modified with a charge tagged photo-cleavable mass marker (**D**).

These probes were twelve bases in length (opposed to fourteen as previously) in order to further destabilise the duplexes containing a single base mismatch, to lower the duplex T_m to allow a lower incubation temperature in the assay.

Initially charged POA monomers **(44)** and **(45)** were synthesised (synthesis discussed in **Chapter 4**, **Scheme 4.8**) to introduce the preformed positive charge essential for the detection of the mass marker cleaved from the PNA probe. Due to the poor coupling efficiencies of these monomers an alternative method of introducing a preformed positive charge was investigated.

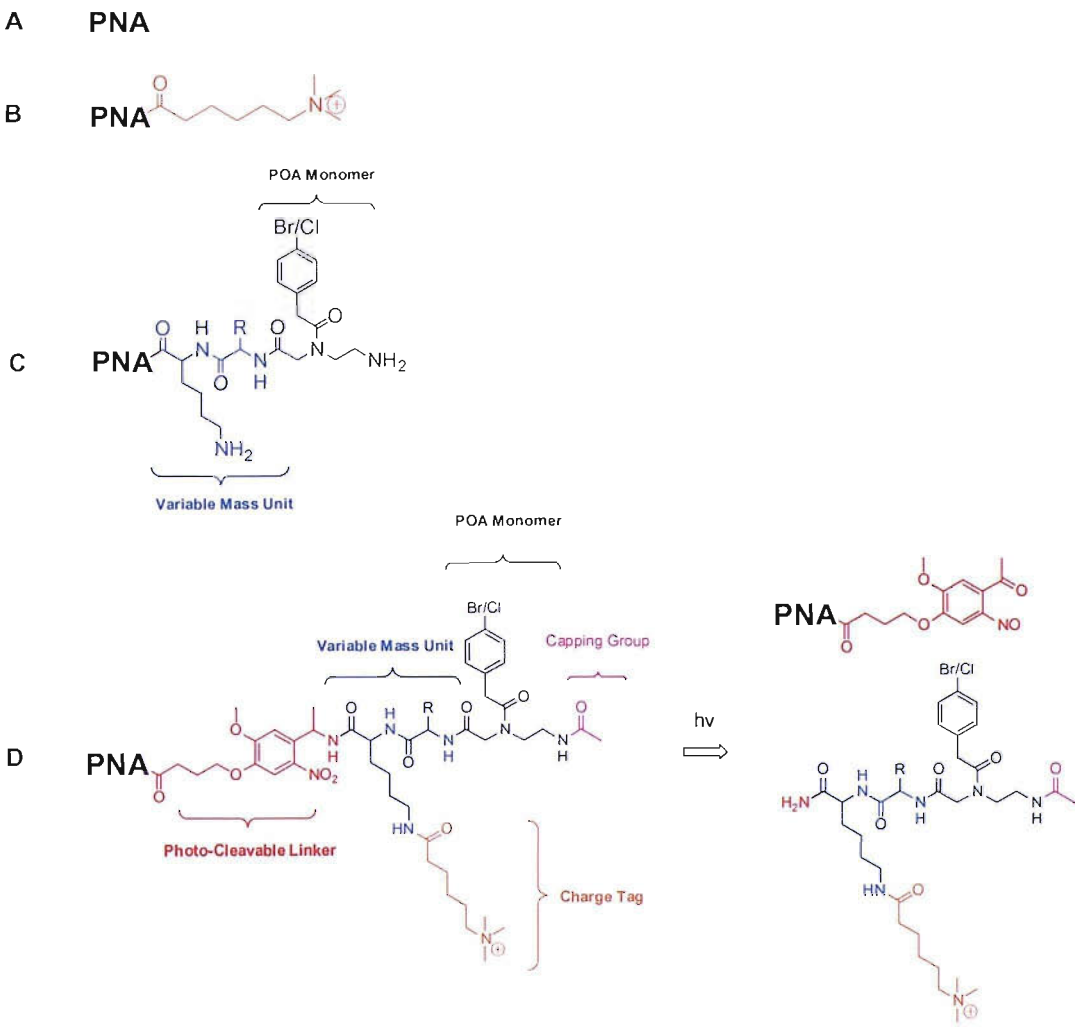


Figure 3.1. The general structure of the PNA probes.

The uncontrolled derivatisation with C_5Q discussed in **Chapter 2** (section 2.2.7.1) is a consequence of having two reactive amino functionalities in the mass marker (amino terminus and lysine side chain). In order to control the derivatisation the amino terminus of the probe was capped with acetic anhydride prior to deprotection from the solid support and removal of the lysine (*tert*-butoxycarbonyl (BOC)) and base protecting groups. The PNA was then cleaved from the resin and all the protecting groups were removed leaving only one reactive amino functionality. The amino functionality on the lysine side chain was selectively derivatised with C_5Q allowing the introduction of a single preformed positive charge on the mass marker (**Figure 3.1**).

PNA I.D.	Sequence	Calculated Average Mass (Da) $(M+H)^+/(M)^+$	Calculated Monoisotopic Mass (Da) $(M+H)^+/(M)^+$
PNA(12**C**)	CTTTCCT**C**CACT	3133.0477	3131.2500
PNA(12T)	CTTTCCTTCACT	3148.0593	3146.2497
PNA^+(12**C**)	($(CH_3)_3{}^+N(CH_2)_5CO$)CTTTCCT**C**CACT	3288.2897	3286.3810
PNA^+(12T)	($(CH_3)_3{}^+N(CH_2)_5CO$)CTTTCCTTCACT	3303.2995	3301.3870
PNA-M^+(12**C**)	(CH_3CO)Br-Ala-($(CH_3)_3{}^+N(CH_2)_5CO$)Lys-CTTTCCT**C**CACT	3826.7296	3823.5397
PNA-M^+(12T)	(CH_3CO)Cl-Phe-($(CH_3)_3{}^+N(CH_2)_5CO$)Lys-CTTTCCTTCACT	3873.3877	3870.6212
PNA-L-M^+(12**C**)	(CH_3CO)Br-Ala-($(CH_3)_3{}^+N(CH_2)_5CO$)Lys-hv-CTTTCCT**C**CACT	4107.0102	4103.6456
PNA-L-M^+(12T)	(CH_3CO)Cl-Phe-($(CH_3)_3{}^+N(CH_2)_5CO$)Lys-hv-CTTTCCTTCACT	4125.6346	4122.7084

Table 3.1. PNA probes prepared to detect the cystic fibrosis W1282X mutation.

3.2. UV-Melting

In order to investigate the ability of the modified and unmodified PNA probes to discriminate between fully complementary and single base mismatched DNA targets, thermal denaturation experiments were carried out with the PNA probes and synthetic DNA fragments (**Table 3.2**). Again, the probes showed excellent discrimination between the perfectly matched and single base mismatched alleles. Hysteresis, a difference between melting and annealing temperatures was observed with the mutant PNA probe PNA(12**T**) and the DNA targets with an **A** and a **T** at the variant site, indicating the system is not in equilibrium (**Figure 3.2**). The melting temperatures of the modified PNA probes were not determined as previous experiments demonstrated that the modifications had little effect on the stability of the PNA/DNA duplex.

PNA I.D.	Oligonucleotide Sequence	Base Pair	T_m (°C)	ΔT_m (°C)
PNA(12C)	ACAGTGTAGGAAAG	C--T	25.2	21.0
PNA(12C)	ACAGTGAAGGAAAG	C--A	ND	ND
PNA(12C)	ACAGTGGAGGAAAG	C--G	46.2	-
PNA(12C)	ACAGTGCAGGAAAG	C--C	ND	ND
PNA(12T)	ACAGTGTAGGAAAG	T--T	29.7	-
PNA(12T) Melting	ACAGTGAAGGAAAG	T--A	57.7	-
PNA(12T) Annealing			49.5	-
PNA(12T) Melting	ACAGTGGAGGAAAG	T--G	45.3	-
PNA(12T) Annealing			35.2	-
PNA(12T)	ACAGTGCAGGAAAG	T--C	25.8	-

Table 3.2. T_m values for the two PNA probes PNA(12C) and PNA(12T) with synthetic DNA targets in 10 mM Tris, 1 M NaCl, pH 7. ΔT_m is the difference in T_m between the fully matched and single base mismatch PNA/DNA duplexes.

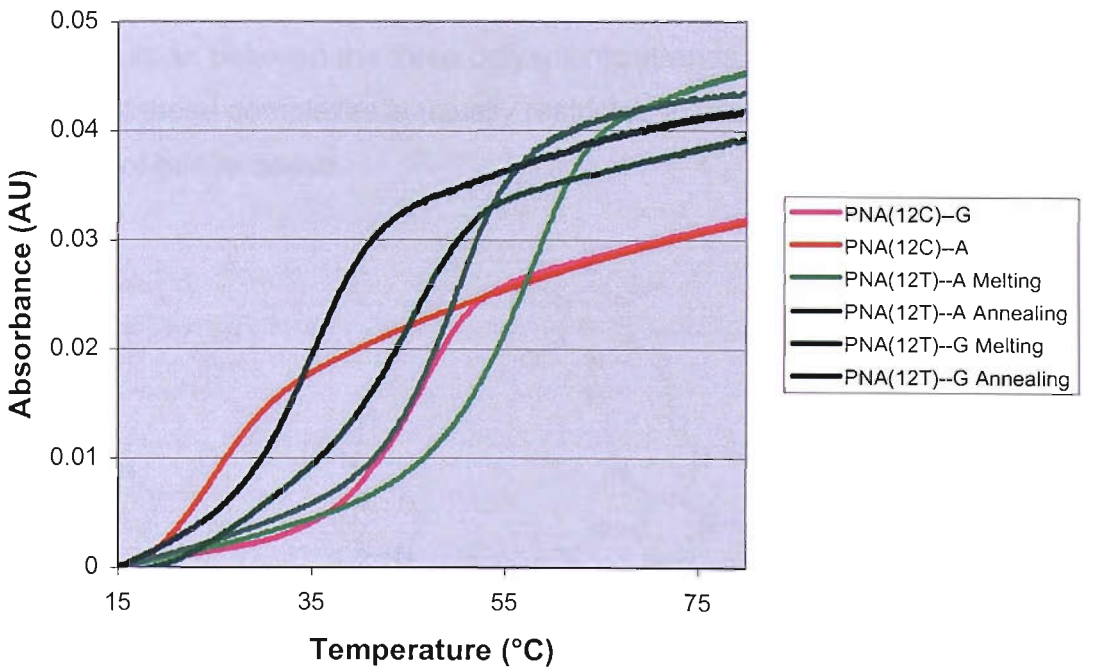


Figure 3.2. UV-Melting profiles of the unmodified PNA probes PNA(12C) and PNA(12T) with synthetic DNA targets in 10 mM Tris, 1 M NaCl, pH 7.

3.3. Triple Helices

DNA triple helices are formed by the binding of synthetic oligonucleotides within the major groove of B-form duplex DNA, where they make specific hydrogen bonds (Hoogsteen) with groups on the exposed faces of the purine bases.[150-152] The third strand can be oriented parallel or antiparallel to the purine strand of the duplex. The parallel motif, where by **T** recognises **A**:**T** and **C** recognises **G**:**C** results in the formation of **C**$^+$.**GC** and **T**.**AT** triplets, which are the most widely studied and are depicted in **Figure 3.3**. Although triple helix formation is highly sequence specific there are a number of reasons for the poor stability of parallel triplexes at physiological pH. Firstly, the **C**$^+$.**GC** triplet is stabilised by protonation and this does not occur at physiological pH (pKa of **C** is 4.5). As a result conditions of low pH are required for protonation of the third strand cytosine. Secondly, triplexes are generally less stable than their duplex counterparts as a consequence of charge repulsion between the three polyanionic strands.[153] In addition, the formation of these complexes is usually restricted to strands primarily composed of purine bases.

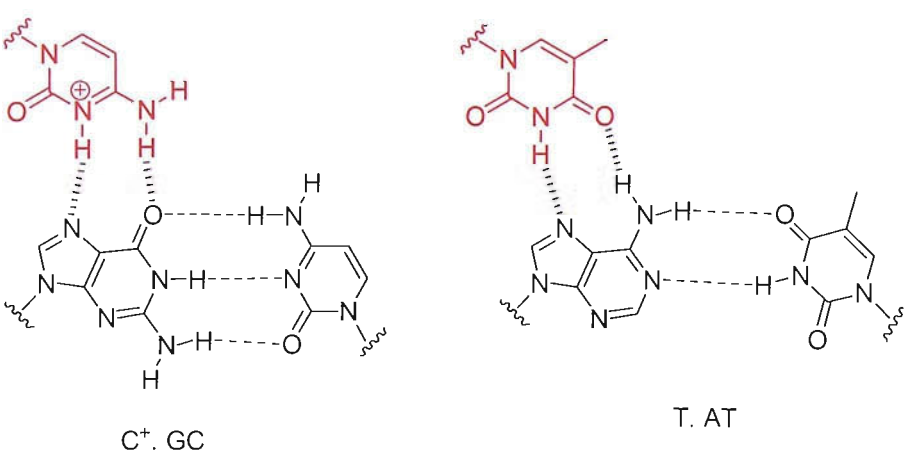


Figure 3.3. Watson Crick (---) and Hoogsteen (ıı) base paring in C$^+$.GC and T.AT triplets.

It is of particular interest that PNAs with a high percentage of pyrimidine bases form stable triplexes with homopurine DNA tracts. The resulting triplexes consist of two PNA strands bound to one DNA strand. Nielsen *et al.* have demonstrated the formation of a $(PNA)_2$/DNA triplex between poly-**T** PNAs and a poly-(d**A**) oligonucleotide in addition to mixed sequence PNAs composed of **T** and **G** and the complementary oligonucleotides composed of **A** and **C**.[119, 154] These complexes have been extensively characterised by linear and circular dichroism (CD) spectroscopy and gel electrophoresis. Two factors contribute to the stability of this type of triple helix. Firstly, the neutral PNA backbone eliminates the strong electrostatic repulsion between the strands of the characteristic $(DNA)_3$ triplex. Secondly, in addition to the Watson Crick and Hoogsteen pairing the $(PNA)_2$/DNA triplex is stabilised by additional hydrogen bonds between the amide nitrogen atoms in the third PNA strand and the oxygen atom in the phosphate of the DNA backbone. These additional hydrogen bonds have been confirmed by X-ray crystallography.[155]

3.4. CD Studies

In order to determine whether the PNA probes were in a triplex in the assay the DNA/PNA interaction was monitored by CD spectroscopy. **Figure 3.4** depicts a CD spectrum of a synthetic DNA target (**ACAGTGGAGGAAAG**). PNA is inherently achiral and exhibits only a weak CD spectrum therefore any changes in the stronger CD spectrum of the DNA target can be attributed to the formation of PNA/DNA complexes.

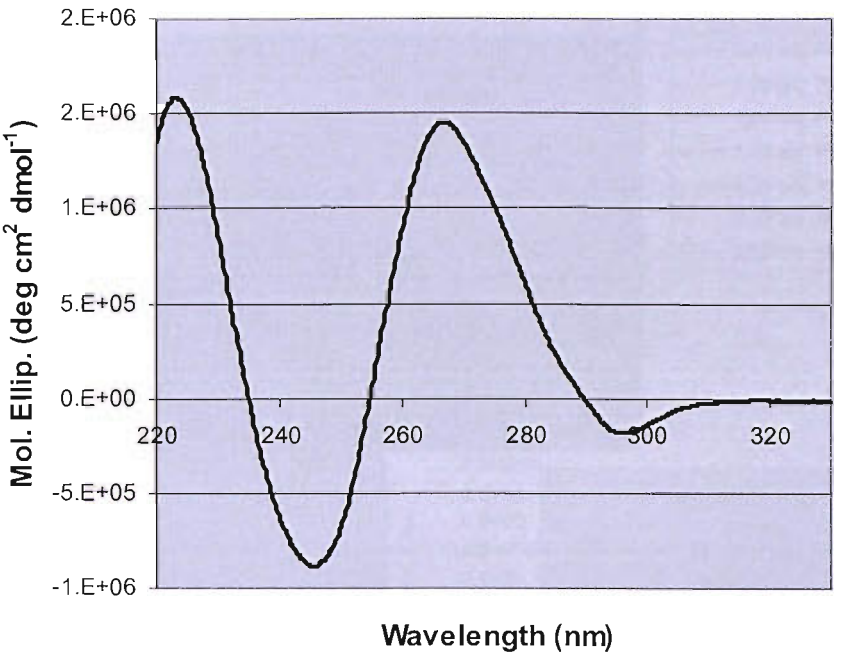


Figure 3.4. CD spectrum of the synthetic DNA target ACAGTGGAGGAAAG.

In order to fully characterise the hybridisation behaviour of the probes with synthetic DNA targets, titration experiments were carried out in the hybridisation buffer. Increasing molar equivalents of the unmodified PNA probes were added to a synthetic DNA target and the CD spectrum was collected. The CD spectrum was accumulated 5 times, smoothed and the spectrum of the buffer was subtracted. A saturation-binding plot was obtained by the subtraction of the molar ellipticity at 265 nm from the molar ellipticity at 276 nm. Analysis of the data shows that the stoichometry of binding for PNA(12**C**) with its complementary DNA target is 1:1 *i.e.* a duplex (**Figure 3.5**). However, the stoichometry of binding for PNA(12**T**) with its complement is 2:1 with titration saturation occurring at 2:1 indicating the formation of a triplex (**Figure 3.6**). These results are not surprising since PNAs were originally designed to bind to duplex DNA in the major groove through Hoogsteen pairing.[117] PNA(12**T**) is predominantly composed of **T** so the triplex formed with its complementary DNA tract is stabilised by the formation of **T.AT** base pairs. In contrast, PNA(12**C**) is mostly composed of **C** and under assay conditions (pH 7) protonation of the third strand cytosine does not occur so no triplex is formed.

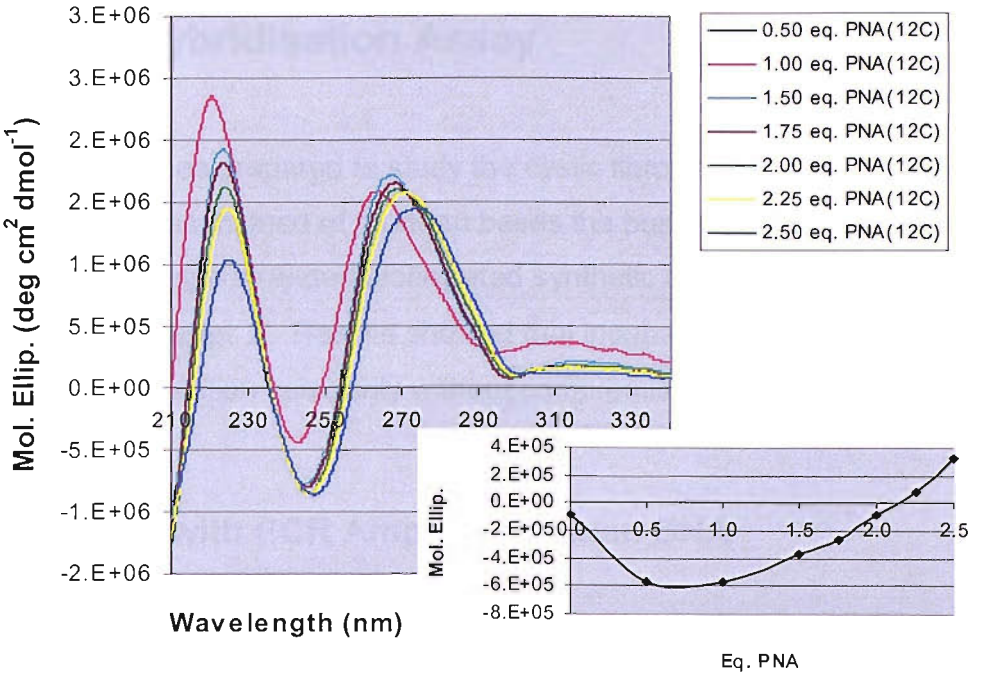


Figure 3.5. CD titration of DNA target (ACAGTGGAGGAAAG) (constant concentration) with PNA(12C). Saturation-binding plot inset.

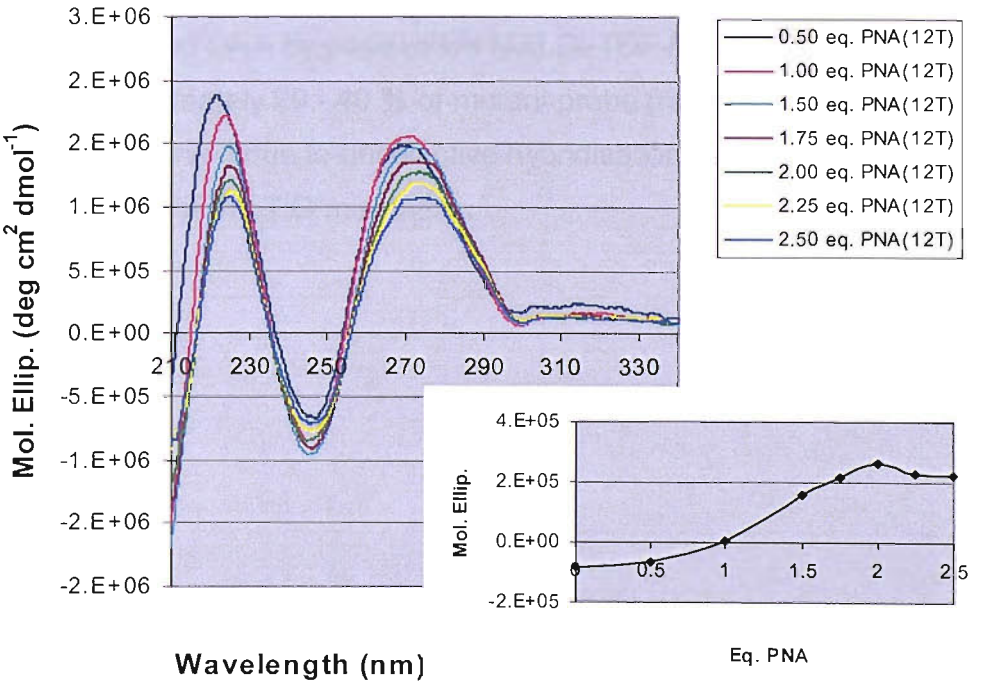


Figure 3.6. CD titration of DNA target (ACAGTGAAGGAAAG) (constant concentration) with PNA(12T). Saturation-binding plot inset.

3.5. PNA Hybridisation Assay

As the PNA probes prepared to study the cystic fibrosis W1282X mutation were twelve bases instead of fourteen bases the basic assay protocol was revised with a single stranded biotinylated synthetic oligonucleotide as described in **Chapter 2**. Results showed that incubation at 55 °C gave adequate hybridisation selectivity without compromising sensitivity.

3.5.1. Assay with PCR Amplified Human DNA

The four different types of probe were used in the PNA hybridisation assay to simulate a G to A mutation in the W1282X locus. In all cases the wild type probe (PNA(12**C**) (**Figure 3.7**), PNA^{+}(12**C**) (**Figure 3.8**), PNA-M^{+}(12**C**) (**Figure 3.9**), PNA-L-M^{+}(12**C**) and (**Figure 3.10**)) were readily detected from wild type amplified DNA by positive ion MALDI-TOF-MS. However, in addition, approximately 20 - 40 % of mutant probe (relative to the wild type probe) was observed due to unselective hybridisation as a consequence of the high stability of the **T**:**G** mismatch.

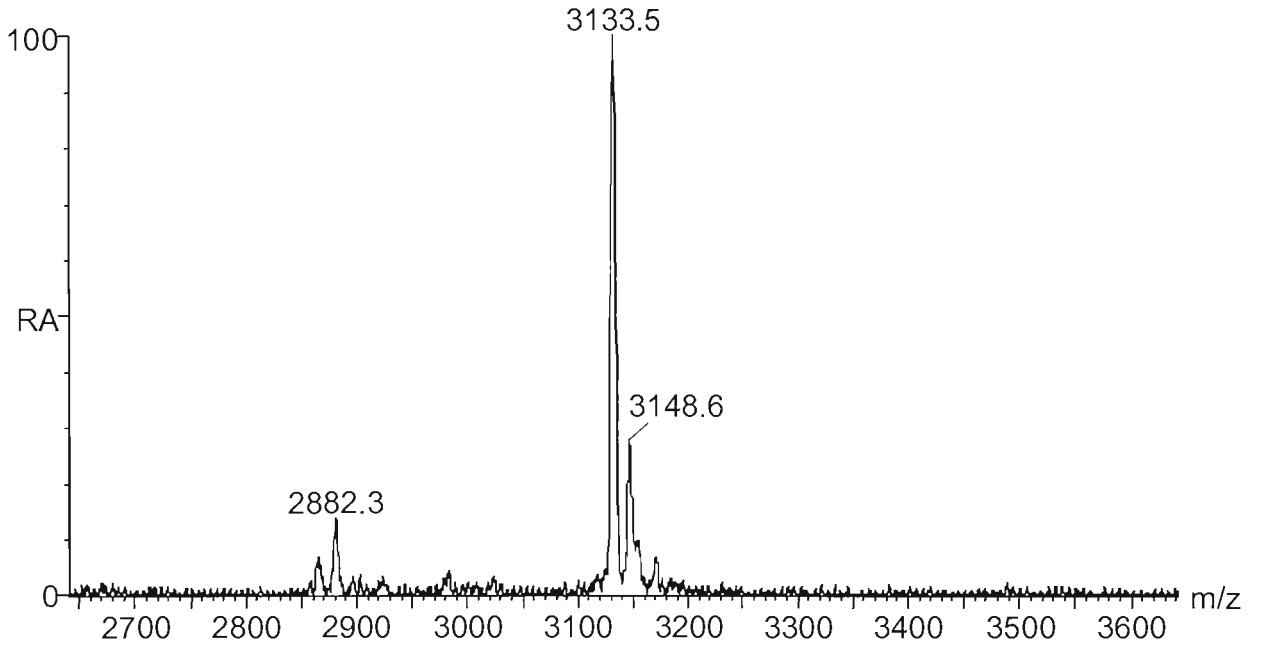


Figure 3.7. Linear MALDI-TOF mass spectrum of the unmodified PNA probe (PNA(12C)) released from the immobilised PCR product. Expected mass $(M+H)^+$ 3133.0 Da (wild type) $(M+H)^+$ 3148.1 Da (mutant).

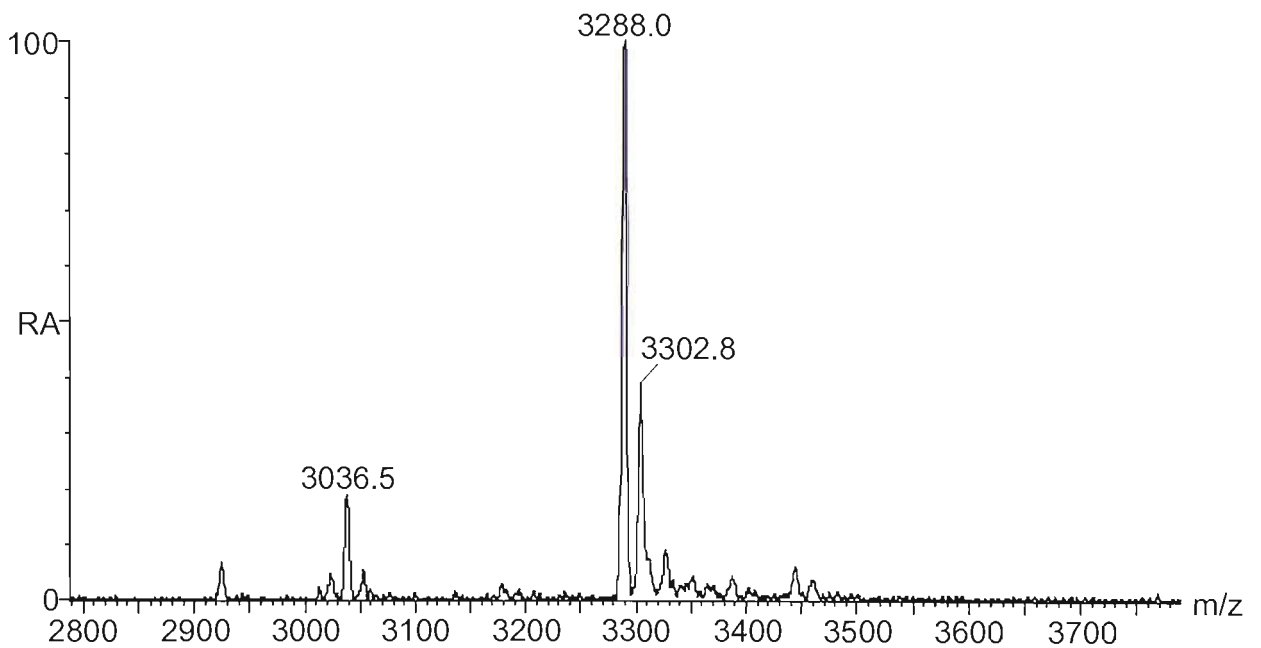


Figure 3.8. Linear MALDI-TOF mass spectrum of the charge tagged PNA probe (PNA^+(12C)) released from the immobilised PCR product. Expected mass $(M)^+$ 3288.3 Da (wild type) $(M+H)^+$ 3303.3 Da (mutant).

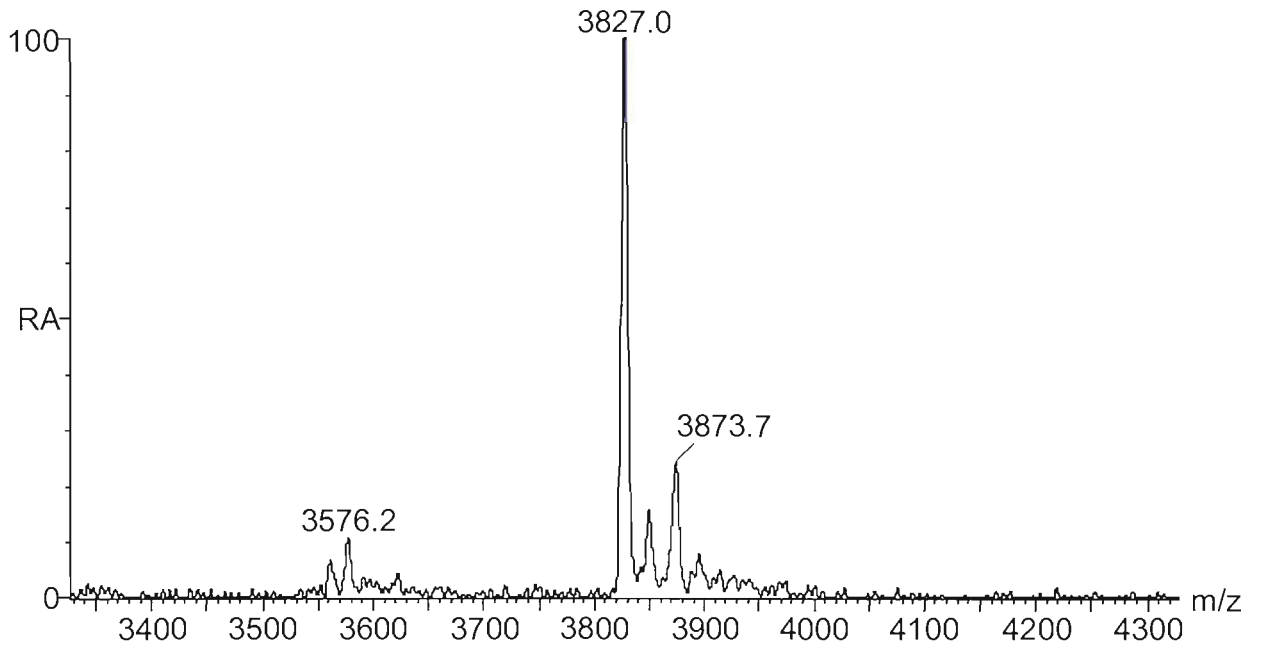


Figure 3.9. Linear MALDI-TOF mass spectrum of the PNA probe modified with charge tagged mass marker (PNA-M^+(12C)) released from the immobilised PCR product. Expected mass $(M+H)^+$ 3826.7 Da (wild type) $(M+H)^+$ 3873.4 Da (mutant).

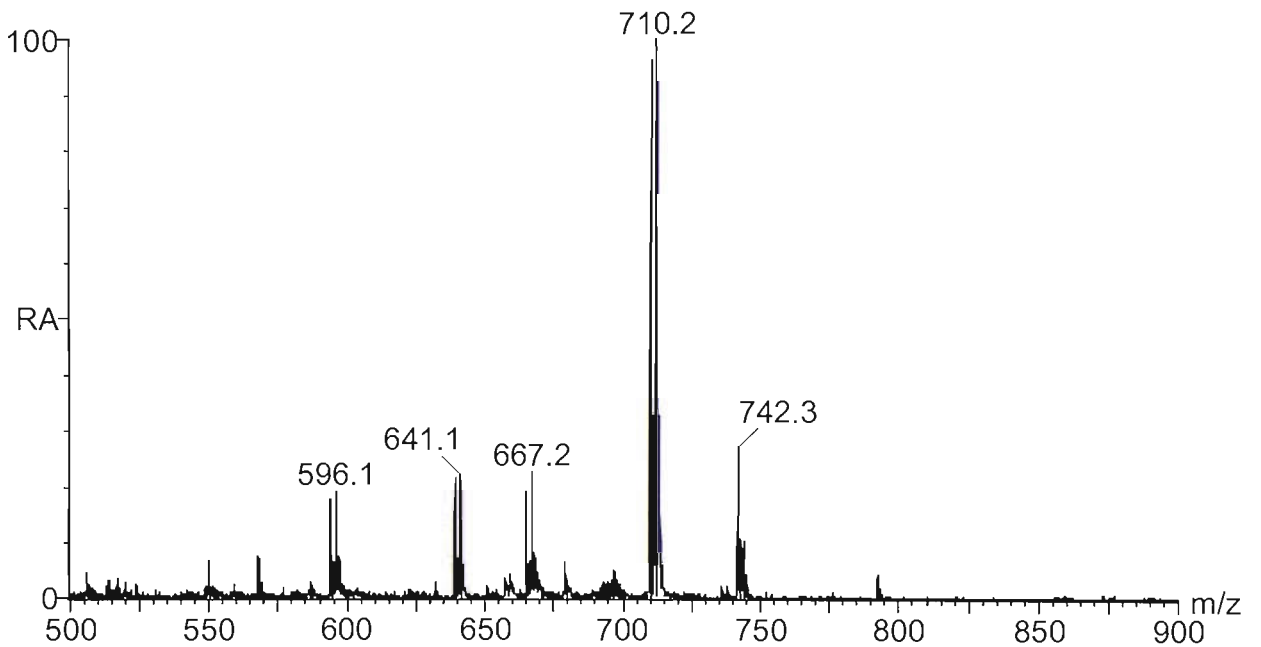


Figure 3.10. Reflectron MALDI-TOF mass spectrum of the PNA probe modified with photo-cleavable charge tagged mass marker (PNA-L-M^+(12C)) released from the immobilised PCR product. Additional ions are observed due to fragmentation of the mass marker due to the high laser poser used to cleave the linker. Expected mass $(M)^+$ 710.3 Da (wild type) $(M+H)^+$ 742.4 Da (mutant).

3.6. Synthesis of PNA Probes

The PNA probes synthesised (PNA(12**C**) and PNA(12**T**)) to detect the **G** to **A** mutation highlighted one of the limitations which is inherent in all hybridisation assays; *i.e.* the difficulty in discriminating between the two probes when stable mismatches occur. In order to alleviate this problem shorter probes were prepared to study a **G** to **T** transversion mutation on the W1282X locus *i.e.* PNA probes with a **C** (wild type) (prepared previously) and **A** (mutant) at the variant site (**Table 3.3**).

PNA I.D.	Sequence	Calculated Average Mass (Da) $(M+H)^+/(M)^+$	Calculated Monoisotopic Mass (Da) $(M+H)^+/(M)^+$
PNA(12**A**)	CTTTCCT**A**CACT	3157.6728	3155.2612
PNA$^+$(12**A**)	($(CH_3)_3{}^+N(CH_2)_5CO$)CTTTCCT**A**CACT	3312.3219	3310.3923
PNA-M$^+$(12**A**)	(CH_3CO)Cl-Phe-($(CH_3)_3{}^+N(CH_2)_5CO$)Lys-CTTTCCT**A**CACT	3882.4012	3879.6328
PNA-L-M$^+$(12**A**)	(CH_3CO)Cl-Phe-($(CH_3)_3{}^+N(CH_2)_5CO$)Lys-hv-CTTTCCT**A**CACT	4162.6817	4159.7387

Table 3.3. PNA probes prepared to detect a T at the variant site.

3.6.1. UV-Melting

In order to identify the optimum incubation temperature and to investigate the ability to discriminate between fully complementary and single base mismatched DNA targets, thermal denaturation experiments were carried out with the unmodified PNA probes with complementary and single base mismatched synthetic DNA targets (**Table 3.4**).

The probes showed excellent discrimination between the perfectly matched and single base mismatched alleles with T_m values differing by up to 15.2 °C between the wild type PNA probe (PNA(12**C**)) and its complement (**G**) and mutant PNA probe (PNA(12**A**)) and the wild type DNA target (**G**).

PNA I.D.	Oligonucleotide Sequence	Base Pair	T_m (°C)	ΔT_m (°C)
PNA(12**A**)	ACAGTG**T**AGGAAAG	**A**--**T**	43.2	-
PNA(12**A**)	ACAGTG**A**AGGAAAG	**A**--**A**	ND	ND
PNA(12**A**)	ACAGTG**G**AGGAAAG	**A**--**G**	31.0	12.2
PNA(12**A**)	ACAGTG**C**AGGAAAG	**A**--**C**	29.0	14.2

Table 3.4. T_m values for the PNA probe PNA(12A) with synthetic DNA targets in 10 mM Tris, 1 M NaCl, pH 7. ΔT_m is the difference in T_m between the fully matched and single base mismatch PNA/DNA duplexes.

3.6.2. Limit of Determination

Assay conditions were reevaluated as the T_m of the **A**:**G** (29.0 °C) mismatch between the wild type probe and mutant DNA target is significantly lower than that of the **T**:**G** (45.3 °C melting and 29.0 °C annealing) mismatch. The incubation temperature at which there is discrimination between the fully matched probe and the probe with a single base mismatch without compromising sensitivity was determined as described previously. Results showed that incubation at 30 °C gave excellent hybridisation selectivity. The sensitivity of the assay was evaluated with single stranded synthetic DNA targets as described previously. The limits of determination for the assay using the unmodified PNA probes (PNA(12**C**) and (PNA(12**A**)) was 100 fmol of single stranded DNA target.

It is of particular interest that a ten-fold increase in sensitivity was observed when using the PNA probes modified with preformed positive charges (charged tagged PNA probes (PNA^+(12**C**) and (PNA^+(12**A**)) and PNA probes modified with charged tagged mass markers (PNA-M^+(12**C**) and (PNA-M^+(12**A**)). The limit of determination for the mass marker, cleaved by the nitrogen laser from the immobilised PNA probe, was 100 fmol of DNA. It is important to note that previous limit of determination experiments with the underivatised mass markers were unsuccessful as the experiments were limited by the quantity of biotinylated DNA target that could be immobilised on the beads.

3.6.3. Assay with PCR Amplified Human DNA

The four different types of probe were used in the PNA hybridisation assay to simulate a **G** to **T** mutation in the W1282X locus. In all cases the wild type probe (PNA(12**C**) (**Figure 3.11**), PNA^+(12**C**) (**Figure 3.12**), PNA-M^+(12**C**) (**Figure 3.13**), PNA-L-M^+(12**C**) and (**Figure 3.14**)) was readily detected from wild type amplified DNA by positive ion MALDI-TOF-MS. In contrast to the previous experiments designed to simulate the **G** to **A** mutation no mutant probe was detected.

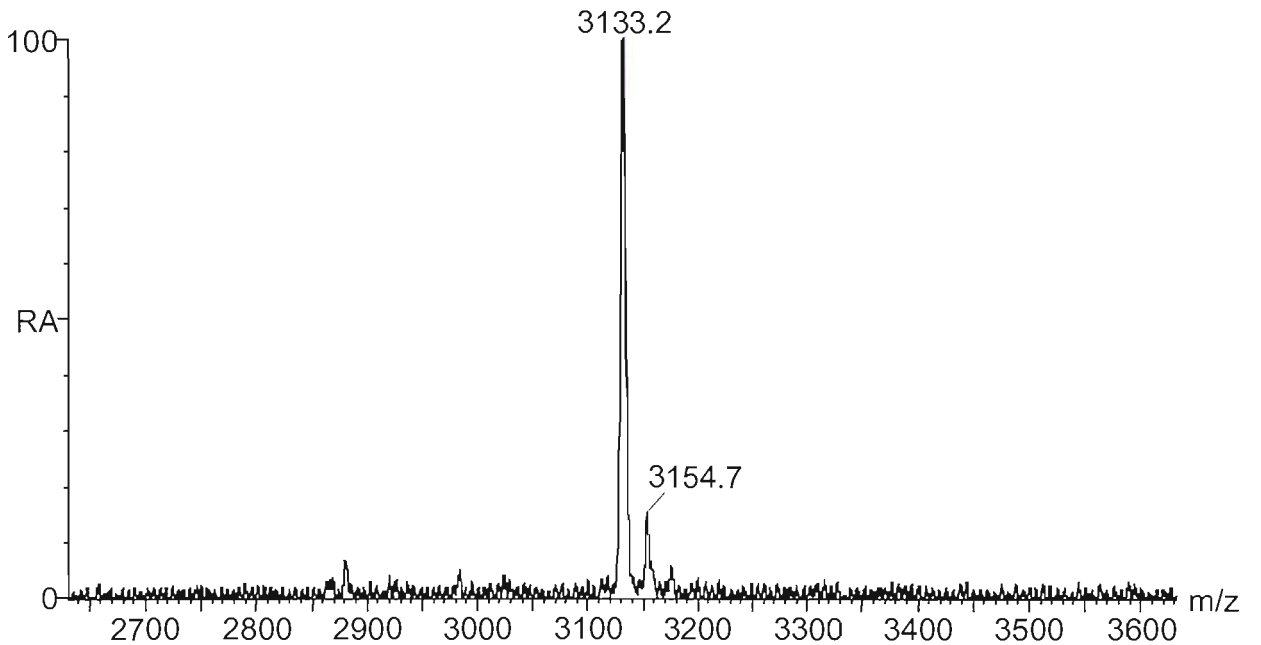


Figure 3.11. Linear MALDI-TOF mass spectrum of the unmodified PNA probe (PNA(12C)) released from the immobilised PCR product. Expected mass $(M+H)^+$ 3133.0 Da.

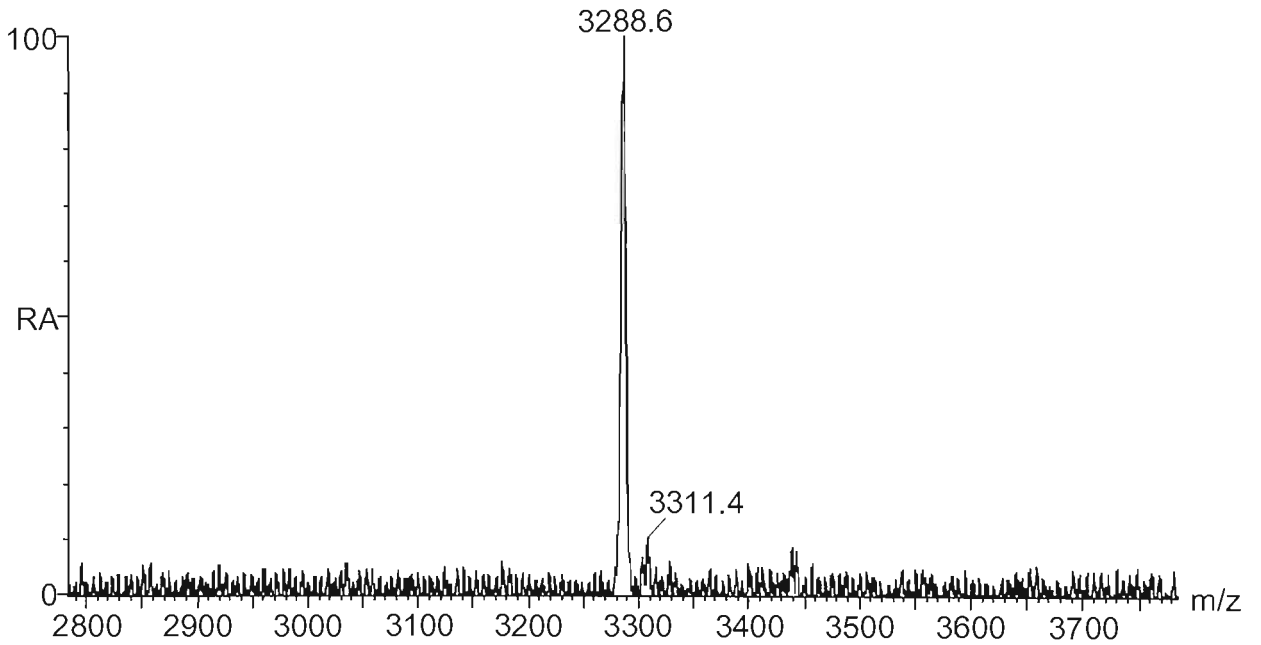


Figure 3.12. Linear MALDI-TOF mass spectrum of the charge tagged PNA probe (PNA^+(12C)) released from the immobilised PCR product. Expected mass $(M)^+$ 3288.3 Da.

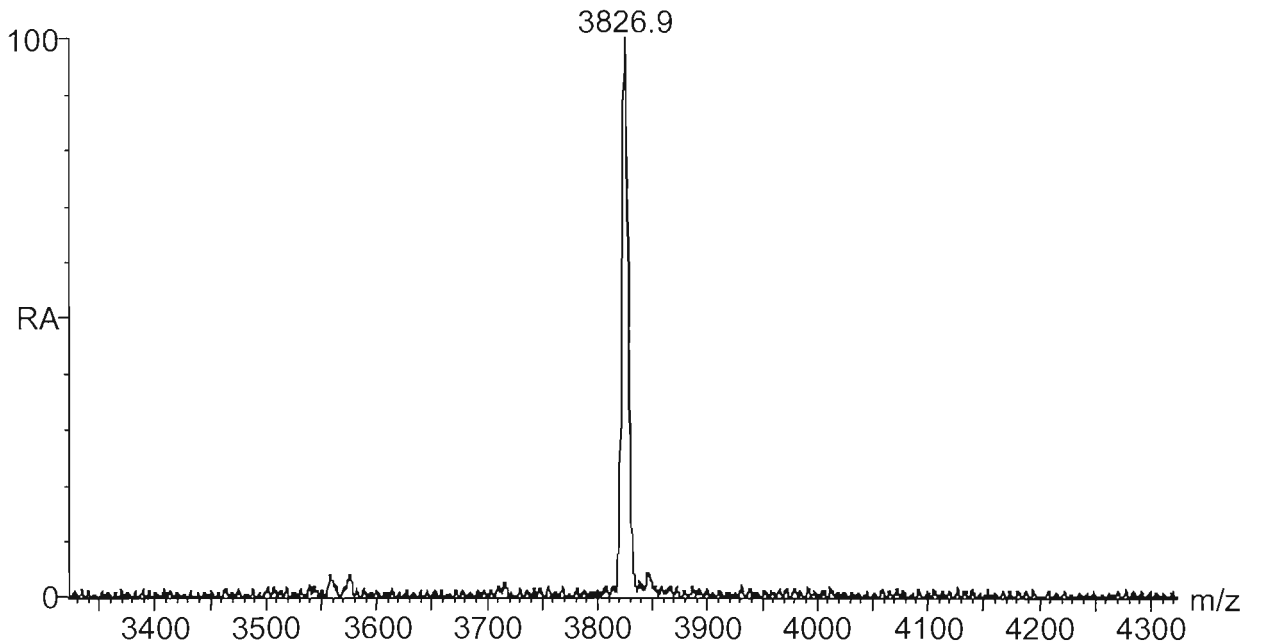


Figure 3.13. Linear MALDI-TOF mass spectrum of the PNA probe modified with charge tagged mass marker (PNA-M^+(12C)) released from the immobilised PCR product. Expected mass $(M+H)^+$ 3826.7 Da.

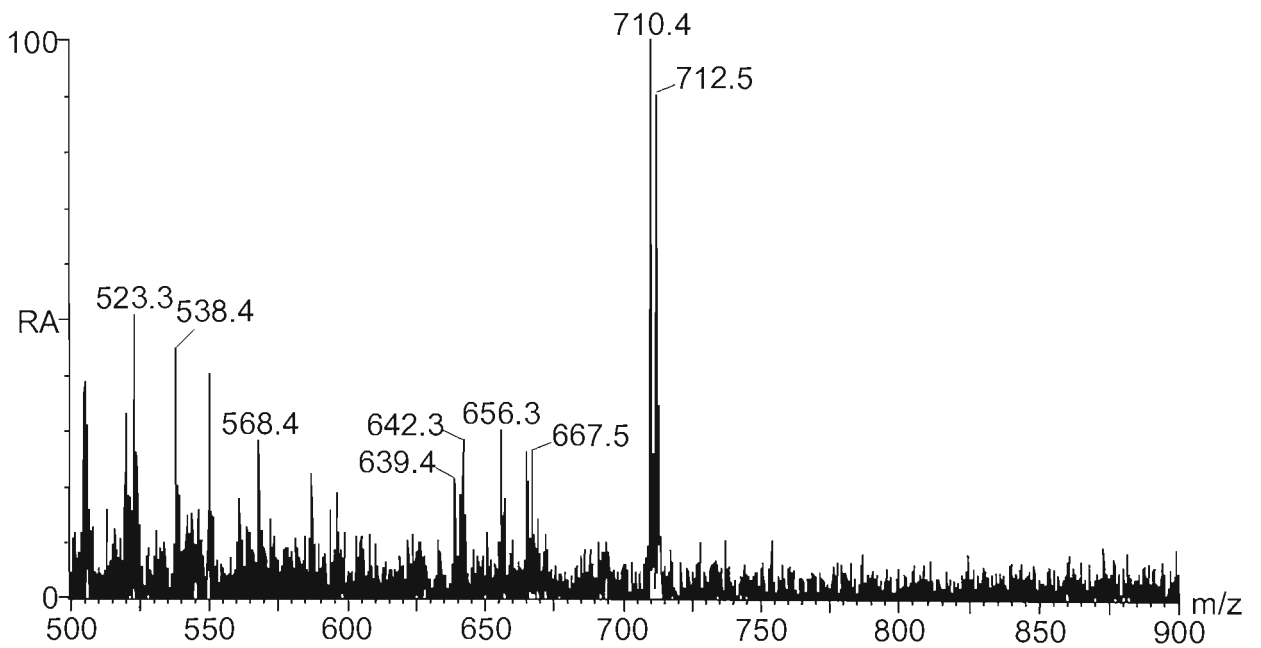


Figure 3.14. Reflectron MALDI-TOF mass spectrum of the PNA probe modified with photo-cleavable charge tagged mass marker (PNA-L-M^+(12C)) released from the immobilised PCR product. Expected mass $(M)^+$ 710.3 Da.

3.7. Transferability of PNA Hybridisation Assay

Throughout this investigation all mass spectra for the hybridisation assay were acquired using a Waters TofSpec2E. To ensure the wide applicability of the assay and to guarantee it could be readily transferred to other MALDI-TOF instruments it was deemed necessary to determine if data could be acquired on MALDI-TOF instruments produced by other manufacturers. To evaluate the compatibility of the hybridisation assay with other instruments, samples were prepared and analysed on the Bruker Ultraflex MALDI-TOF-MS at Syngenta, Jealott's Hill.

3.7.1. Limit of Determination

The limits of determination for the assay using the unmodified PNA probes (PNA(12C) and (PNA(12A)) was 1 pmol of single stranded DNA target and 100 fmol of DNA for all other types of probe. Whilst a ten-fold decrease in the limit of determination was observed it is probable that sensitivity of the assay on the Bruker instrument is comparable to that on the Waters TofSpec2E for a number of reasons. Firstly, the samples were prepared and subsequently analysed 48 hours later. It is possible that some degradation of the sample had occurred prior to analysis leading to a decrease in sensitivity. In addition, all data on the Ultraflex MALDI-TOF-MS was acquired in reflectron mode. It would be expected that greater sensitivity would be achieved in linear mode however this would also have resulted in a loss of mass resolution. These results confirm the ten-fold increase in sensitivity observed when using the probes derivatised with a preformed positive charge and highlight a comparable sensitivity between the two instruments, which in turn, suggests the applicability of the assay to other MALDI-TOF instruments.

3.7.2. Assay with PCR Amplified Human DNA

MALDI-TOF data were successfully obtained for the three different types of PNA probe ((PNA(12**C**) (**Figure 3.15**), PNA-M^+(12**C**) (**Figure 3.16**) and PNA-L-M^+(12**C**) (**Figure 3.17**)). In addition, there were no significant differences in the results obtained on the TofSpec2E and the Ultraflex indicating the method could be readily transferred to other MALDI-TOF instruments. The Ultraflex is a newer instrument and has been constructed to higher instrument specifications with respect to mass accuracy and mass resolution. As a result better mass resolution has been achieved and it is easier to distinguish between the unmodified PNA probes, which in turn should allow a greater degree of multiplexing.

It is worth noting that a greater than expected error in mass measurement was observed. This was attributed to the difference in laser powers used when acquireing the sample and calibration data. The use of a high laser power is not ideal but is essential in the case of the probe with the photo-cleavable mass marker for efficient cleavage of the linker.

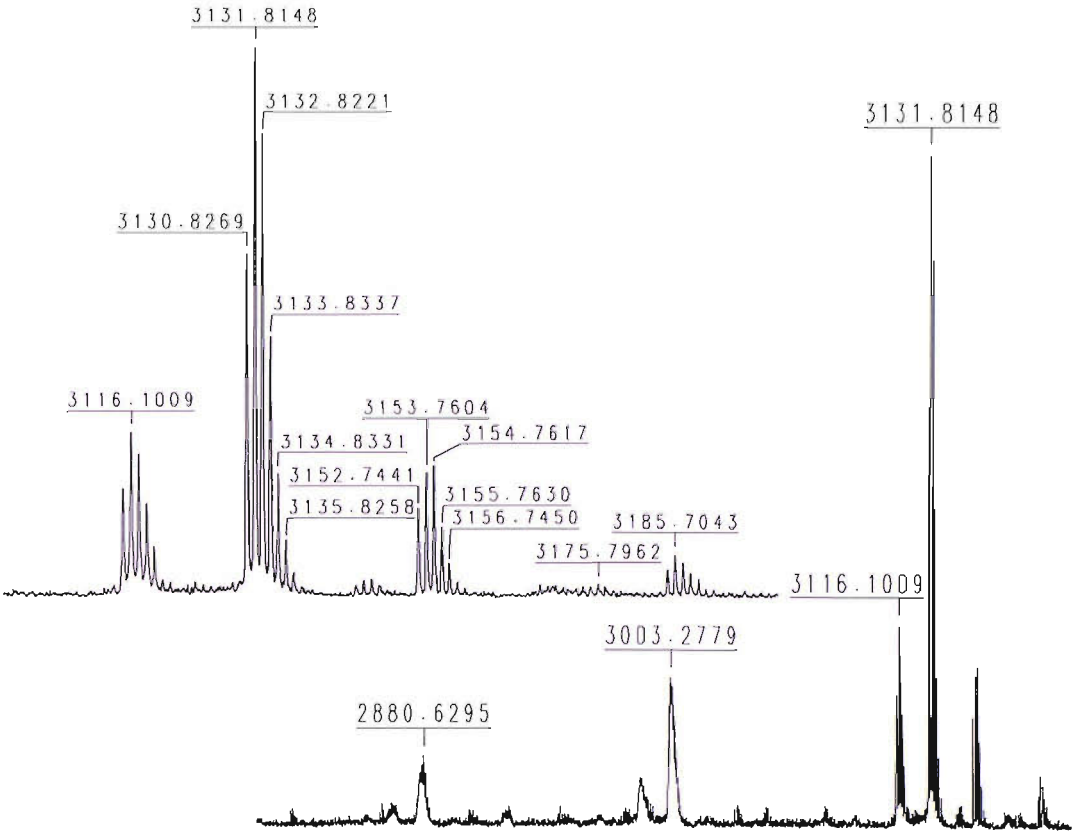


Figure 3.15. Reflectron MALDI-TOF mass spectrum obtained on the Brucker Ultraflex of the unmodified PNA probe (PNA(12C)) released from the immobilised PCR product (externally calibrated). Expected mass $(M+H)^+$ 3131.3 Da.

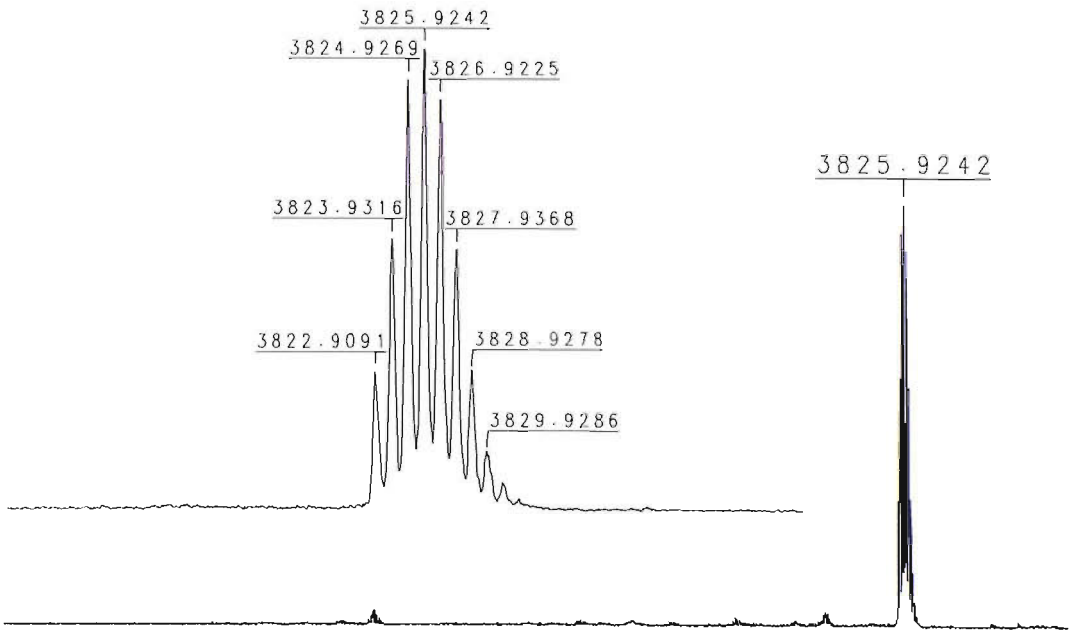


Figure 3.16. Reflectron MALDI-TOF mass spectrum obtained on the Bruker Ultraflex of the unmodified PNA probe (PNA-M^+(12C)) released from the immobilised PCR product (externally calibrated). Expected mass $(M)^+$ 3823.54 Da.

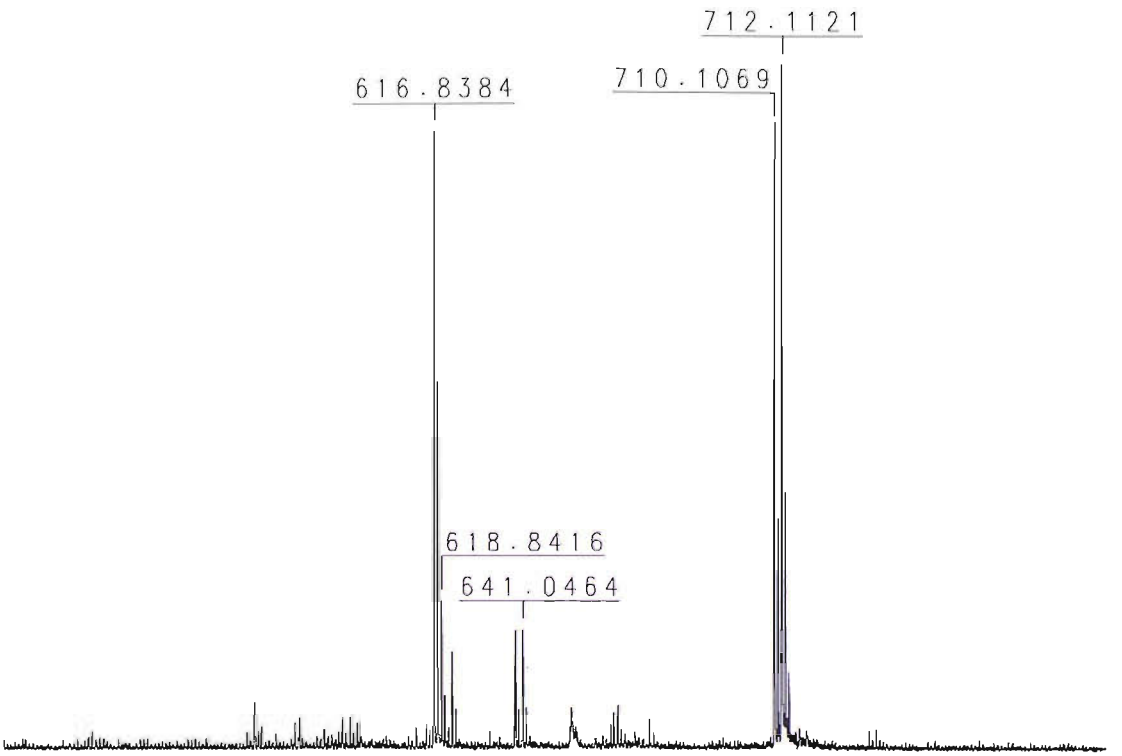


Figure 3.17. Reflectron MALDI-TOF mass spectrum of the PNA probe modified with photo-cleavable charge tagged mass marker (PNA-L-M^+(12C)) released from the immobilised PCR product (externally calibrated). Expected mass $(M)^+$ 710.3 Da.

3.8. Summary

Four different types of PNA probe have been prepared for detection of the W1282X cystic fibrosis mutation. The synthesis of the probes was modified to include the introduction of a preformed positive charge. All types of probes have been tested in the hybridisation assay with wild type PCR amplified DNA however these experiments have highlighted two drawbacks of this method. Firstly, a limitation associated with all hybridisation assays, stable mismatches introduce a problem as full discrimination cannot always be achieved between the fully complementary duplexes and those containing a single base mismatch. In addition, PNAs composed predominately of pyrimidine have been shown to form triplexes with homopurine DNA targets. Whist the formation of triplexes in the assay does not directly impose a problem it is not ideal and sequences which are likely to form triplexes should be avoided in future studies.

Limit of determination experiments with probes designed to detect a **G** to **T** mutation have shown the introduction of the preformed positive charge gives a ten-fold increase in the sensitivity of the assay over the unmodified probe. In addition, the preformed positive charge allowed the limit of determination for the mass marker, cleaved by the nitrogen laser from the immobilised PNA probe to be evaluated.

The assay was evaluated on a Bruker Ultraflex MALDI-TOF mass spectrometer to ensure the assay could be readily transferred to other MALDI-TOF instruments. The results showed compatibility of the hybridisation assay with other instruments and in addition, showed a comparable sensitivity.

4. Chemical Synthesis

4.1. Background

Developed by Nielsen and co-workers in 1991, PNA is an analogue of DNA in which the phosphodiester backbone has been replaced with a synthetic peptide composed of repeating *N*-(2-aminoethyl) glycine units (**Figure 2.1**).[116, 117] Individual nucleobases are attached to each of the units *via* methylene carbonyl linkers to provide a molecular framework, which enables PNA oligomers to hybridise to their complementary nucleic acids obeying the Watson Crick base pairing rules.[118, 119]

The methods adopted for the assembly of PNA are very similar to those used in the preparation of peptides, which can be performed in solution or on a solid support. Solid phase synthesis is more commonly employed as it has distinct advantages over solution based methods in particular; speed, high yields and the ease of work up and purification. The carboxy terminus of the PNA monomer is anchored to an insoluble resin and an excess of the subsequent residue is added in the presence of a coupling reagent. An amide bond is made between the amino functionality of the immobilised monomer and the carboxy terminus of the incoming residue. The protecting group on the amino functionality of the incoming residue remains unaffected during the coupling of the monomer and is removed at the end of each cycle (**Figure 4.1**). This process is repeated allowing extension of the PNA oligomer. Reactive functionalities are protected with groups which are stable to the reagents used during the synthesis and are removed at the end of the synthesis during the final cleavage from the solid support.

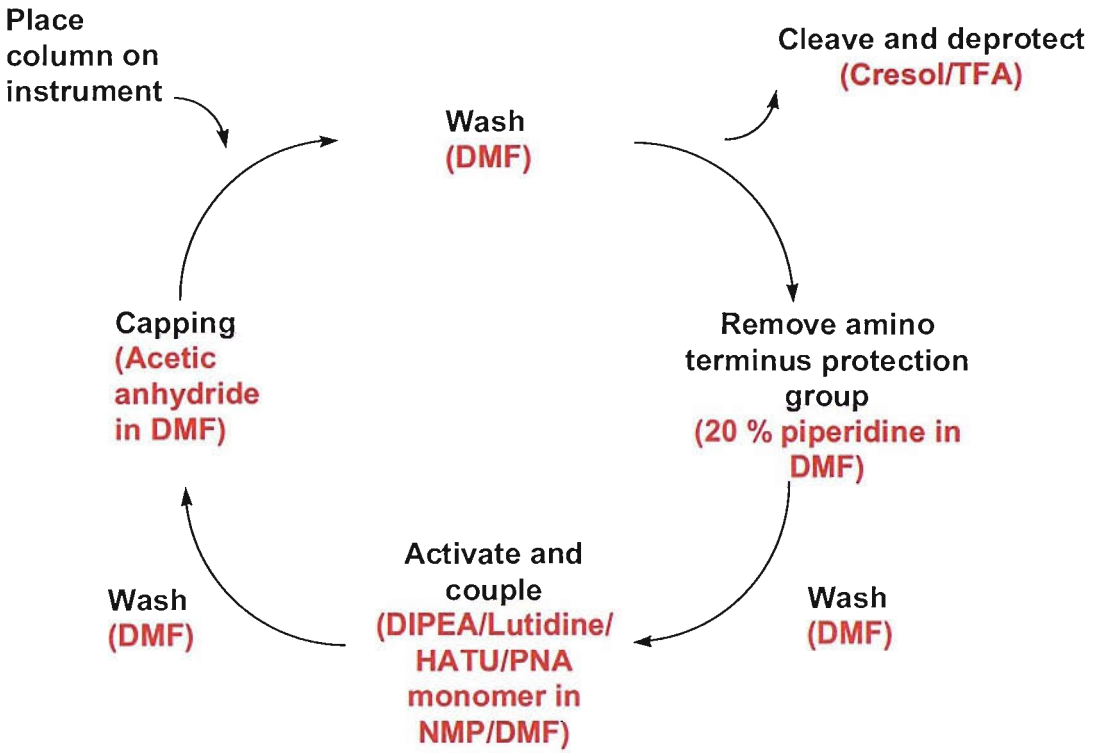


Figure 4.1. The PNA synthesis cycle. The reagents used for the addition of Fmoc/bhoc PMA monomers are highlighted in red.

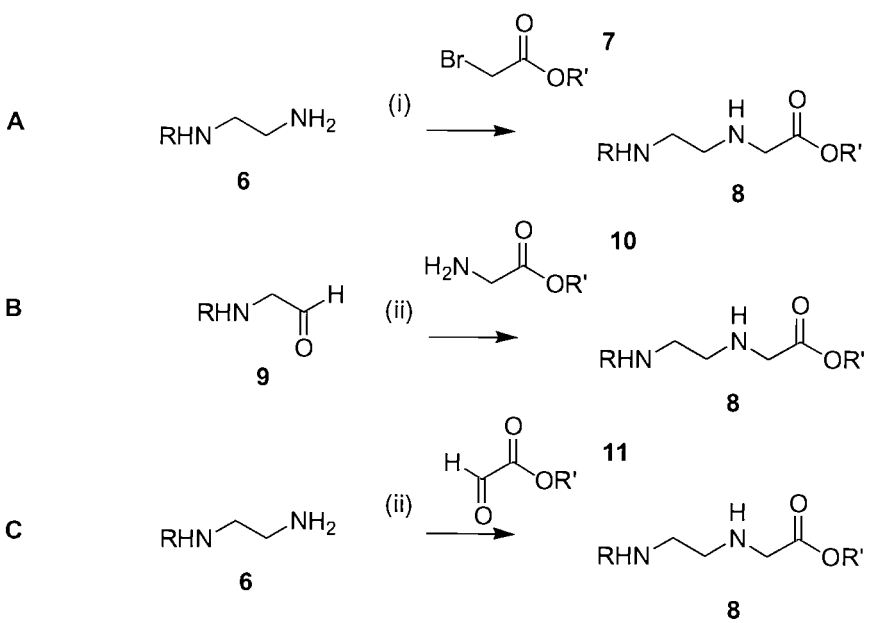
This stepwise process allows the modification of the PNA oligomers prior to release from the solid support by any compound with an amine reactive functionality such amino acids or non base pairing monomers or fluorescent reporter groups.[156-158]

The chemistry of PNA synthesis, like the chemistry of peptide synthesis is based on the choice of compatible protecting groups for the reactive moieties. The approach employed is usually characterised by the choice of the amino protecting group. Whilst many different strategies have been employed for the protection of the amino functionality the acid labile Boc and base labile Fmoc are among the more popular choices.[126, 159]

The synthesis of the PNA monomers can be divided into two sections, the preparation of the backbone **(22)** and the preparation of the nucleobase acetic acids (*e.g.* **(51)** and **(55)**). The two moieties are then efficiently coupled together using standard peptide coupling reagents such as *N*-hydroxybenzotriazole (HOBt) and *N*-ethyl-*N*'-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDC).[126, 160]

The assembly of the PNA backbone has been accomplished using several different protocols. The following are some of the most frequently employed:

- Nucleophilic substitution of the *mono*-protected ethylene diamine (**Scheme 4.1.A**).[160-162]
- Reductive amination of the *N*-protected amino acetaldehyde with glycine esters (**Scheme 4.1.B**).[163]
- Reductive amination of the glyoxylic esters with the mono-protected ethylene diamine (**Scheme 4.1.C**).[164]



Reagents and conditions i) DMF or CH_2Cl_2 rt. ii) $NaCNBH_3/CH_3CO_2H$, or H_2/Pd, CH_3OH, rt.

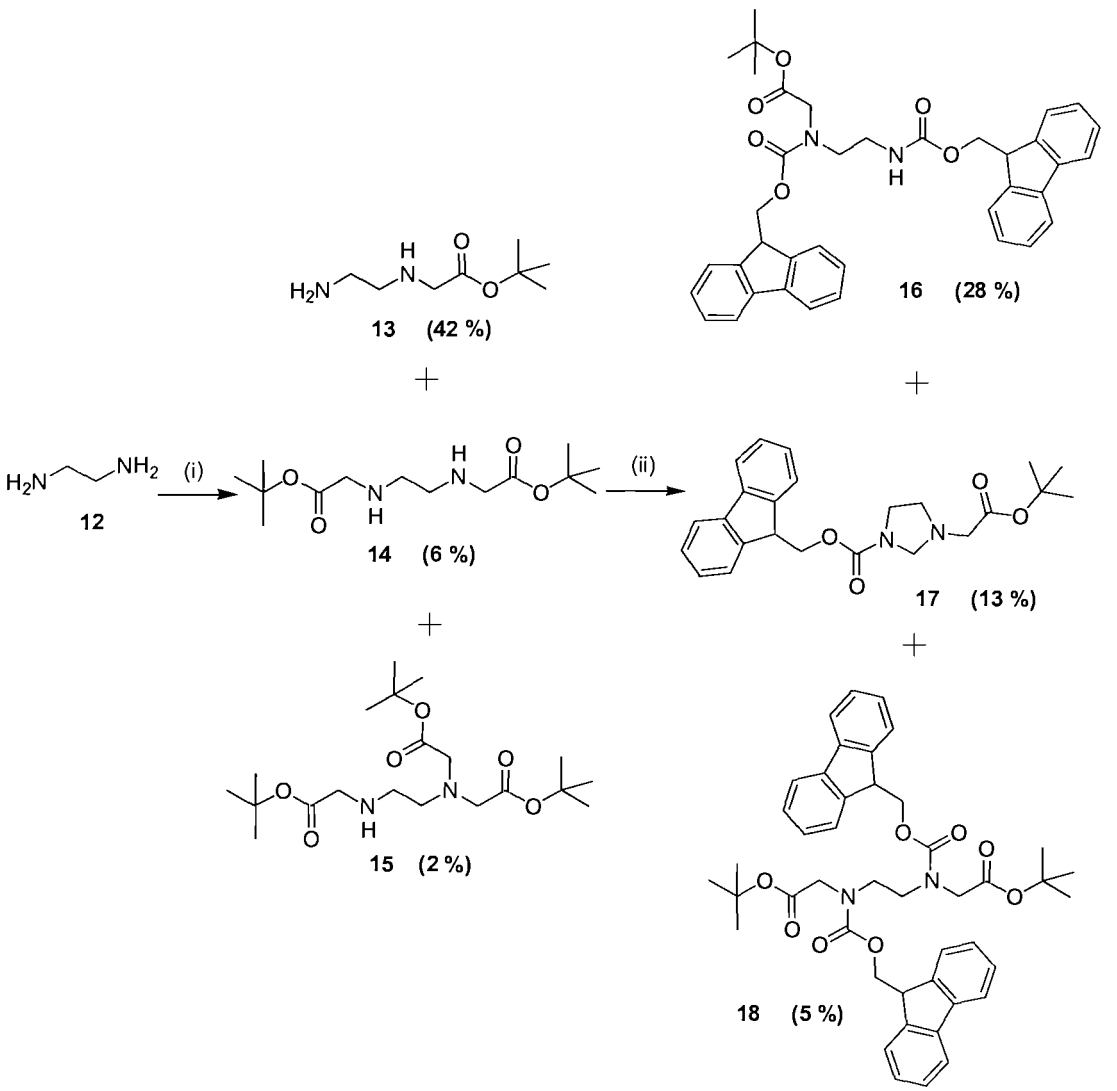
Scheme 4.1. Synthesis of PNA backbone.

Thymine can be readily converted to the base acetic acid by alkylation with a bromoacetic acid ester. However, the exocyclic amines of adenine, cytosine, and guanine must be protected to avoid undesirable side reactions during oligomer synthesis and to aid the solubility of the resulting monomer.[160] In most cases the base is protected with an orthogonal protecting group to the *N*-terminal amine to allow selective removal of each group.

While the Boc/ benzyloxycarbonyl (CBZ) protection strategy is commonly used, Fmoc protected monomers offer many advantages, namely milder deprotection conditions and improved monomer stability.[160] For these reasons the Fmoc strategy was used for PNA synthesis. The carboxy terminus of the PNA backbone could then be temporarily masked as its *tert*-butyl ester to be removed after coupling to the nucleobase acetic acid, enabling assembly all PNA oligomers on the solid support using Fmoc chemistry.

4.2. Synthesis of the Fmoc Protected PNA Backbone (22)

The synthesis of the initial target, the Fmoc protected PNA backbone **(22)** was attempted (**Scheme 4.2**), following the procedure described by Thompson *et al.*[160] Alkylation of excess ethylene diamine **(12)** at 0 °C with *tert*-butyl bromoacetate followed by an aqueous workup to remove any residual ethylene diamine yielded a mixture of products, the *mono*-alkylated amine **(13)** and the *di*- and *tri*- substituted products **(14)** and **(15)**. Again, the reaction yielded a number of products, which were isolated by column chromatography and characterised by NMR and mass spectrometry. It is probable that imidazolidine **(17)** is formed by reaction with dichloromethane whilst compounds **(16)** and **(18)** are the products of uncontrolled Fmoc protection. This series of reactions was repeated a number of times, but could not be used to generate any of the target.

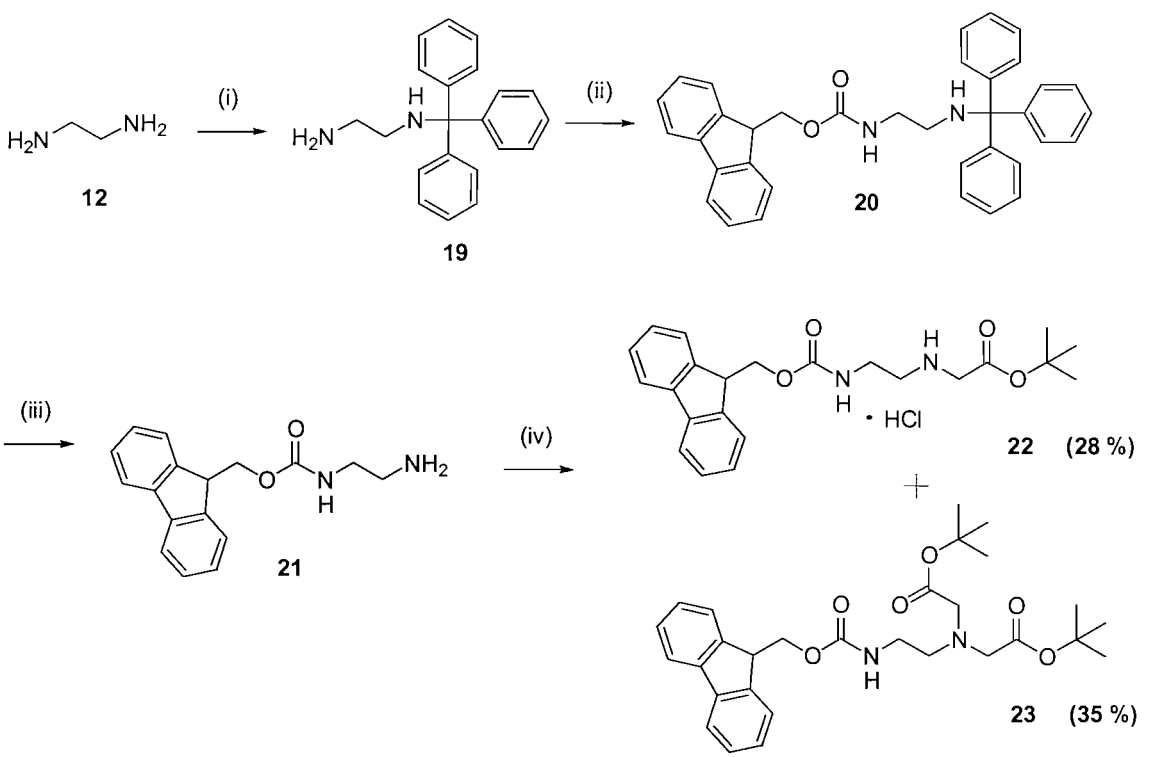


Reagents and conditions i) *tert*-Butyl bromoacetate (0.1 eq), CH_2Cl_2, 0 °C, 15 h, **(13)** 42 %, **(15)** 6 %, **(16)** 2 %. ii) Fmoc-Su (1.0 eq), DIPEA (1.0 eq), CH_2Cl_2, 0 °C, 45 min, **(16)** 28 %, **(17)** 13 %, **(18)** 5 %.

Scheme 4.2. Attempted synthesis of PNA backbone (22) according the procedure described by Thompson and co-workers.[160]

An alternative approach was devised (**Scheme 4.3**) in which ethylene diamine was selectively capped with a trityl-protecting group. Tritylation of excess ethylene diamine **(12)** followed by aqueous workup gave amine **(19)** in excellent yield. This enabled selective Fmoc protection with Fmoc-Su to generate the *di*-protected product **(20)**.

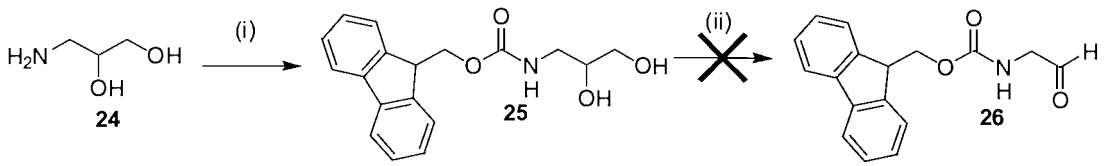
The trityl moiety was removed by treatment with trifluoroacetic acid (TFA) in the presence of a cation scavenger (triethylsilane). Tritutation of the product with ether removed any residual organic impurities yielding primary amine **(21)** as a colorless solid. In the final step, an excess of amine **(21)** was alkylated with *tert*-butyl bromoacetate and the desired product was isolated by column chromatography. Despite using an excess of amine **(21)**, which was recovered from the crude mixture by precipitation from ethyl acetate and hexane, it was not possible to control the alkylation reaction and the *di*-alkylated compound **(23)** was always the dominant product. Although this method was not subjected to a comprehensive evaluation of solvents and reaction optimisation, it was not considered suitable for the synthesis of the required quantities of PNA backbone **(22)**.



Reagents and conditions i) Trityl chloride (0.2 eq), pyridine, rt, 4 h, 97 %. ii) Fmoc-Su (1.0 eq), DIPEA (1.0 eq), CH_2Cl_2, 0 °C, 1 h, 70 %. iii) TFA, triethyl silane, CH_2Cl_2, rt, 30 min, 100 %. iv) *tert*-Butyl bromoacetate (1.0 eq), DIPEA (1.0 eq), DMF, rt, 1 h, **(22)** 28 %, **(23)** 35 %.

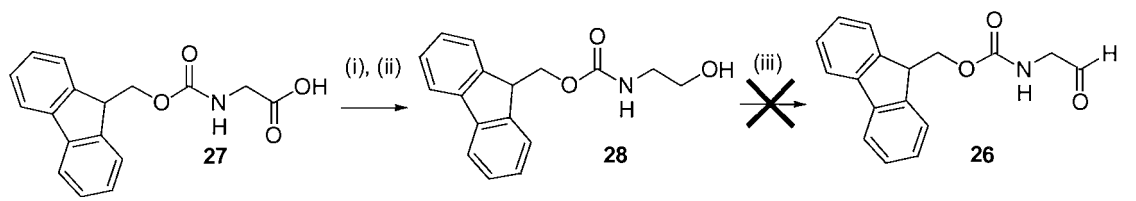
Scheme 4.3. Synthesis of PNA backbone (22).

Salvi and co-workers have described the synthesis of Fmoc-pseudo peptides by reductive amination of Fmoc-amino aldehydes with sodium cyanoborohydride.[163] In an attempt to prepare the Fmoc-protected PNA backbone **(22)** *via* the reductive amination of protected amino acetaldehyde with a glycine ester, efforts were made to synthesise the Fmoc-protected amino acetaldehyde **(26)** for coupling with *tert*-butyl 2-aminoacetate. The two different approaches employed to assemble the target aldehyde **(26)** are outlined in **Scheme 4.4** and **4.5**.



Reagents and conditions i) Fmoc-Su (1.0 eq), CH_2Cl_2, rt, 12 h, 61 %. ii) $NaIO_4$ (2.0 eq), CH_2Cl_2 or $NaIO_4$ (2.0 eq), CH_3OH/H_2O or $NaIO_4$ (2.0 eq), Acetone/H_2O.

Scheme 4.4. Synthesis of the Fmoc-protected amino acetaldehyde (26).

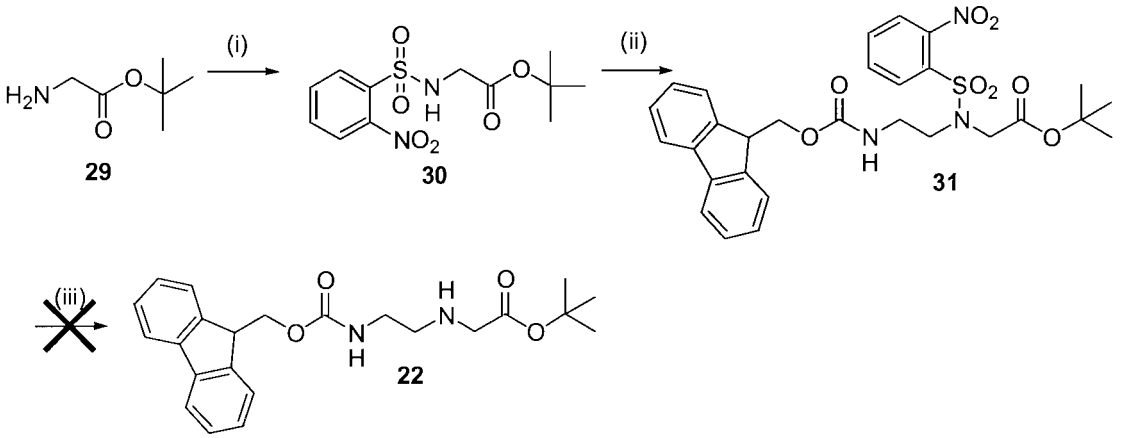


Reagents and conditions i) *iso*-Butyl chloroformate (1.0 eq), NMM (1.0 eq), DME, - 15 °C, 10 min. ii) $NaBH_4$ (1.5 eq), H_2O, - 15 °C, 10 min, 85 % (over 2 steps). iii) TPAP (5.0 mol %), NMO (1.5 eq), powdered molecular sieves, CH_2Cl_2 or TPAP (5.0 mol %), NMO (1.5 eq), powdered molecular sieves, CH_2Cl_2/CH_3CN or TPAP (5.0 mol %), NMO (1.5 eq), powdered molecular sieves, CH_3CN.

Scheme 4.5. Synthesis of the Fmoc-protected amino acetaldehyde (26).

Protection of 3-amino-1,2-propanediol **(24)** under basic conditions with Fmoc-Su gave diol **(25)** in 61 % yield. Cleavage of diol **(25)** with excess sodium periodate was attempted in various solvents (water, dichloromethane, methanol/water, acetone/water). However no mixture of the solvents investigated dissolved sufficient quantities of both the diol **(25)** and sodium periodate for the reduction to be successful.[165] One further attempt to prepare Fmoc-protected amino acetaldehyde **(26)** was made. Formation of the mixed anhydride of acid **(27)** in the presence of 4-methylmorpholine (NMM) and isobutyl chloroformate and subsequent reduction with aqueous sodium borohydride yielded alcohol **(28)** in good yield. All endeavors to oxidise alcohol **(28)** to aldehyde **(26)** with 4-methylmorpholine *N*-oxide (NMO) and catalytic quantities of tetrapropylammonium perruthenate (TPAP) in various solvents (dichloromethane, acetonitrile, dichloromethane/acetonitrile) failed.[166]

Fukuyama and co-workers have described an efficient synthesis of secondary amines from nitrobenzenesulfonamides.[167] Modification of the reported method enabled the preparation of sulfonamide **(30)** under basic conditions from 2-nitrobenzenesulfonyl chloride (NsCl) and amine **(29)** (**Scheme 4.6**). Sulfonamide **(30)** was efficiently coupled to alcohol **(28)** (prepared as described in **Scheme 4.5**) under standard Mitsunobu conditions to yield sulfonamide **(31)**. Treatment of sulfonamide **(31)** with one equivalent of 1-benzenethiol and three equivalents of potassium carbonate or one equivalent of 1-benzenethiol and one equivalent of DIPEA led to removal of both the sulfonyl moiety and the Fmoc protecting group.

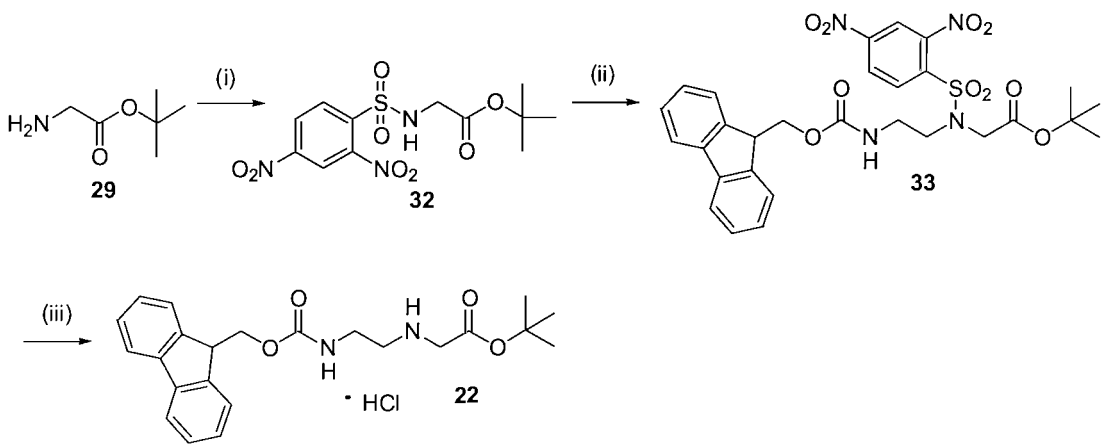


Reagents and conditions i) NsCl (1.0 eq), Et_3N, CH_2Cl_2, rt, 1 h, 85 %. ii) DEAD (1.0 eq), PPh_3 (1.5 eq), Alcohol **(28)** (0.7 eq), CH_2Cl_2, rt, 30 min, 71 %. iii) PhSH (1.2 eq), K_2CO_3 (3.0 eq), CH_3CN or $HSCH_2CO_2H$ (1.3 eq), Et_3N (2.0 eq), CH_2Cl_2.

Scheme 4.6. Synthesis of PNA backbone (22) according the procedure described by Fukuyama and co-workers.[167]

Fukuyama and co-workers published a subsequent report using 2,4–dinitrobenzenesulfonamides in which the sulfonyl moiety was removed under milder conditions (mercaptoacetic acid and DIPEA).[168] Following this procedure sulfonamide **(32)** was efficiently alkylated under Mitsunobu conditions to yield the *N*,*N*-di-substituted 2,4-dinitrobenzene sulfonamide **(33)**. The deprotection of **(33)** to give amine **(22)** as the hydrochloride salt was achieved in high yields by treatment with excess mercaptoacetic acid in the presence of DIPEA.

Using a method adapted from that described by Fukuyama *et al.* amine **(22)** has been prepared in 58 % overall yield as described in **Scheme 4.7**.[168] It is particularly advantageous that this synthetic route requires minimal chromatographic separation and despite being longer than the alternative procedure described by Thompson *et al.*[160] in our hands this route has proven to be a reliable and high yielding synthesis.



Reagents and conditions i) DNsCl (1.0 eq), lutidine, CH_2Cl_2, - 15 °C, 1 h, 91 %. ii) DIAD (1.5 eq), PPh_3 (1.5 eq), Alcohol **(28)** (1.0 eq), CH_2Cl_2, rt, 1 h, 82 %. iii) $HSCH_2CO_2H$ (2.0 eq), DIPEA (3.0 eq), CH_2Cl_2, rt, 1 h, 92 %.

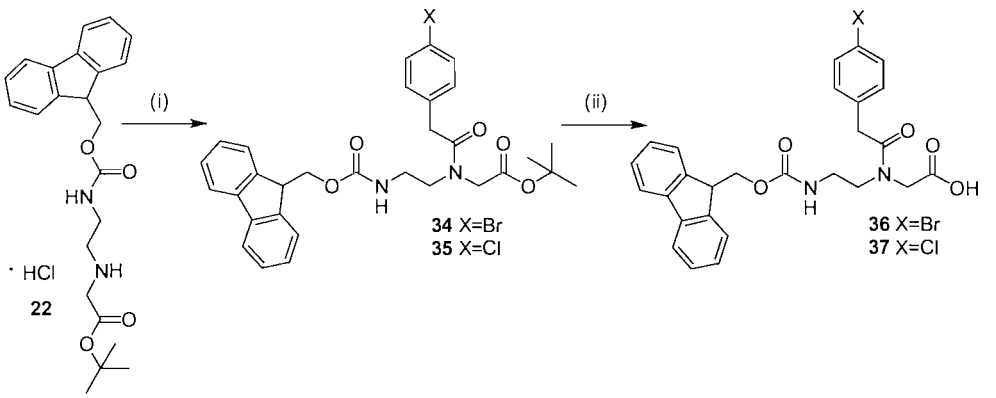
Scheme 4.7. Synthesis of PNA backbone (22) according the procedure described by Fukuyama *et al.*[168]

4.3. Synthesis of the Peptide Organic Acid Monomers

The substitution of the nucleobases of PNA with other moieties such as pyrene, anthraquinone and phenyl has been exploited to produce PNA-intercalator conjugates.[156-158] The monomers were readily prepared by coupling commercially available acetic acid derivatives to the Fmoc-protected PNA backbone and the PNA oligomers were assembled *via* solution phase or automated solid phase synthesis. The aromatic residues have been shown to be capable of stacking but not hydrogen bonding and have subsequently been used as probes to understand the thermodynamic factors involved in PNA binding to DNA and the influence of the individual nucleobases on the overall structure of the hybrid PNA/DNA double helix.[158] As these types of monomers could be readily prepared, they were chosen for incorporation into the PNA probes in conjunction with two amino acids to provide each PNA probe sequence with a unique molecular weight.

The POA monomers were isotopically labelled with either a bromine or chlorine for facile identification of the low molecular weight mass marker once cleaved from the PNA probe.

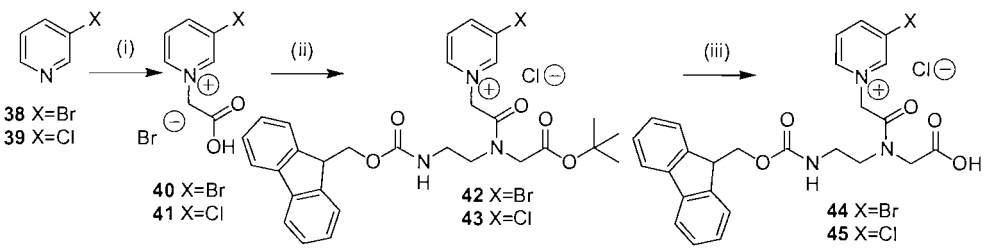
With the Fmoc-protected PNA backbone in hand the POA monomers **(36)** and **(37)** were assembled as described in **Scheme 4.8**. Reaction of the appropriate commercially available phenylacetic acid with backbone **(22)** in the presence of EDC afforded the protected monomers **(34)** and **(35)** in moderate yields. Hydrolysis of the ester with TFA proceeded to give the desired POA monomers **(36)** and **(37)** in near quantitative yields.



Reagents and conditions i) 4-Bromophenyl acetic acid or 4-Chlorophenyl acetic acid (1.1 eq), EDC (1.5 eq), DIPEA, DMF, 2 h, **(34)** 54 %, **(35)** 67 %. ii) TFA:CH_2Cl_2 (1:1), rt, 1h, **(36)** 95 %, **(37)** 96 %.

Scheme 4.8. Synthesis of the Fmoc-protected POA monomers (36) and (37).

Recently Barry and co-workers have described the synthesis of novel pyridinium compounds for transformation of poorly or non-ionisable analytes into compounds readily detected by ES-MS.[169] It was thought that 2-(3-Bromo-1-pyridiniumyl)acetic acid bromide salt **(38)** and the chlorinated analogue could be readily coupled to the PNA backbone **(22)** to make charge tagged mass markers to overcome the derivatisation problem observed with C_5Q. The synthesis of the bromo and chloro charged tagged monomers **(44)** and **(45)** is described in **Scheme 4.9**. The appropriate halogenated pyridine was alkylated in good yields with bromoacetic acid. The pyridinium acetic acid was coupled to the PNA backbone in moderate yield and the *tert*-butyl protecting group was then removed with TFA to yield the charged tagged POA monomers. Attempts to prepare PNA oligomers containing the charged tagged POA monomers were unsuccessful. This was due to the poor coupling efficiencies of the charged tagged POA monomers.

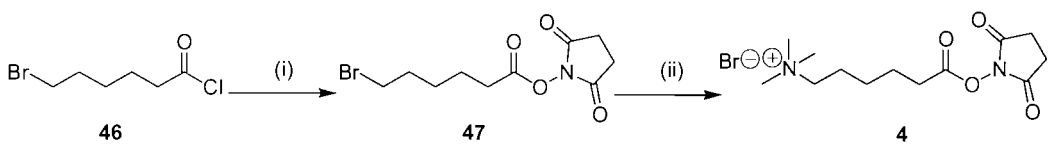


Reagents and conditions i) Bromoacetic acid (1.0 eq), CH_3CN, 90 °C, 2 h, **(40)** 90 %, **(41)** 86 %. ii) PNA backbone **(22)** (0.9 eq), EDC (1.3 eq), DIPEA (1.0 eq) DMF, rt, 1 h, **(42)** 43 %, **(43)** 58 %. iii) TFA:CH_2Cl_2 (1:1), rt, 1 h, **(44)** 96 %, **(45)** 92 %.

Scheme 4.9. Synthesis of the Fmoc-protected charged tagged POA monomers (44) and (45).

4.4. Synthesis of 6-[(2,5-Dioxotetrahydro-1*H*-1-pyrrolyl)oxy]-6-oxohexyl(trimethyl)ammonium bromide (4)

Bartlet-Jones *et al.*[145] have described a method of attaching a quaternary ammonium charge tag to the amino terminus of peptides for increased sensitivity. This method has subsequently been used by Gut and co-workers for charge tagging oligonucleotides *via* a primary amino aliphatic group.[103, 105] The charge tag C_5Q, was synthesised in high yields according to the procedure described by Bartlet-Jones *et al.* (**Scheme 4.10**).[145]



Reagents and conditions i) *N*-hydroxysuccinimide (1.1 eq), CH_3CN, 90 °C, 12 h, 76 %. ii) Trimethylamine (3.0 eq), THF, rt, 12 h, 97 %.

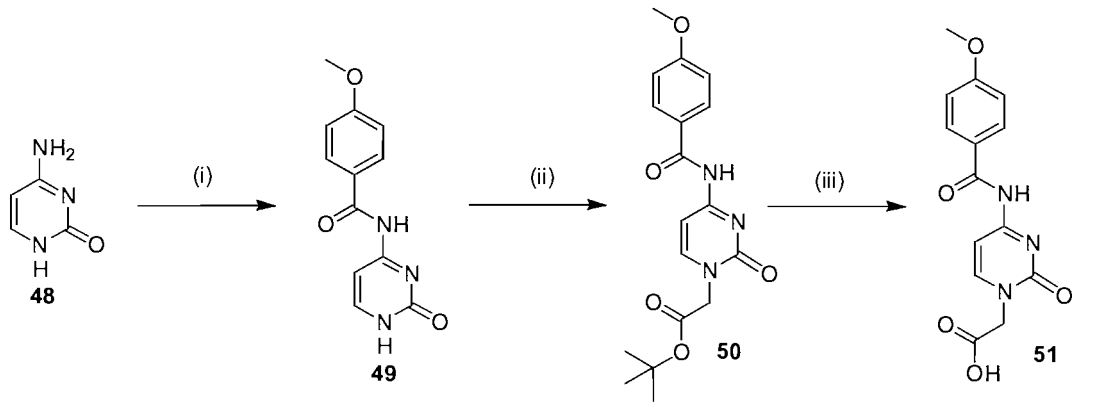
Scheme 4.10. Synthesis of 6-[(2,5-dioxotetrahydro-1*H*-1-pyrrolyl)oxy]-6-oxohexyl(trimethyl)ammonium bromide (4) (C_5Q) according the procedure described by Bartlet-Jones *et al.*[145]

4.5. Synthesis of PNA Monomers

In order to make the assay amenable to ES-MS an acid labile linker was purchased for incorporation into the PNA probes. As the Rink linker[170] is cleaved in TFA the commercially available Fmoc/benzhydryloxycarbonyl (Bhoc) PNA monomers used in the synthesis of the PNA probes with photo-labile linker were unsuitable. The anisoyl-protecting group is cleaved under strong basic conditions, but is stable to the conditions used to remove Fmoc protecting groups so was identified as a suitable group for protection of the PNA bases. The Fmoc/anisoyl cytosine and adenine PNA monomers were synthesised following a combination of previously published procedures.[160, 164, 171, 172]

4.5.1. Synthesis of Nucleobase Acetic Acids

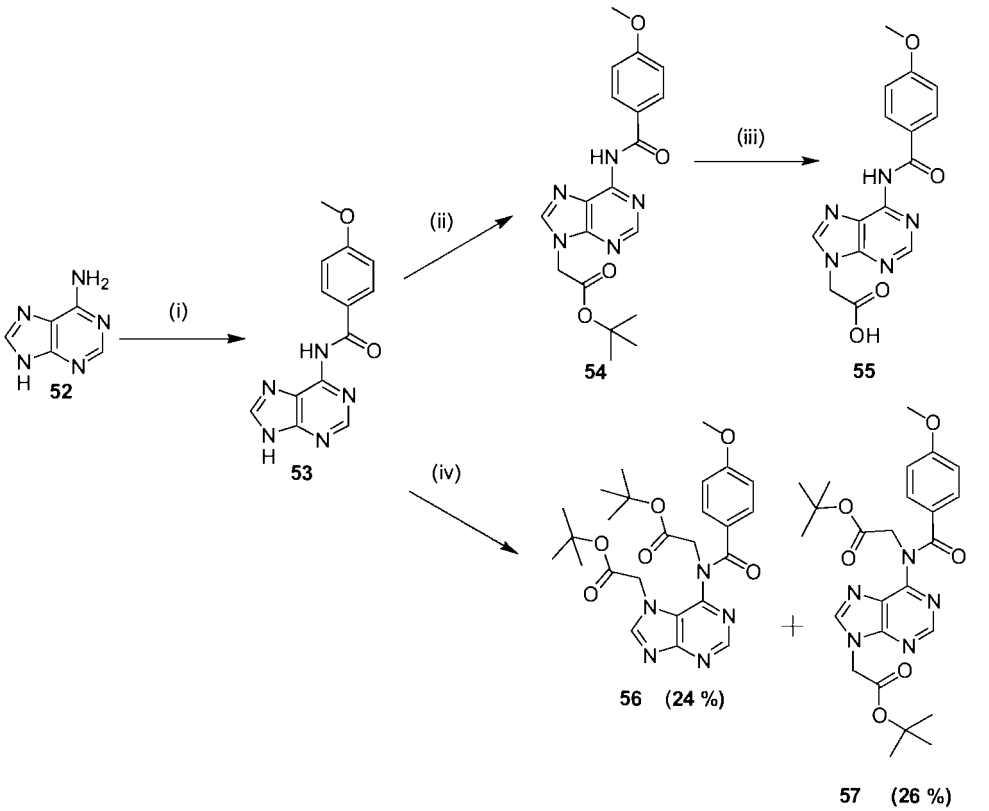
The cytosine acetic acid **(51)** was prepared as described in **Scheme 4.11**. Initially, the 4-amino amine was protected with 4-anisoyl chloride in excellent yield. Due to the poor solubility of compound **(49)**, in common organic solvents, no purification and little characterisation was performed however, analysis by TLC showed minor traces of un-reacted 4-anisoyl chloride. The N1 position was then alkylated with *tert*-butyl bromoacetate affording ester **(50)** in 69 % yield. Hydrolysis of the *tert*-butyl ester with HCl gave acid **(51)** in quantitative yield.



Reagents and conditions i) 4-Anisoyl chloride (1.2 eq), pyridine, rt, 50 min, 89 %. ii) *tert*-Butyl bromoacetate (1.0 eq), DMF, rt, 48 h, 69 %. iii) 4 M HCl, 1,4-dioxane, rt, 12 h, 98 %.

Scheme 4.11. Synthesis of the protected cytosine acetic acid (51).

The protection of the 6-amino of adenine required harsher conditions (**Scheme 4.12**). Adenine was heated to 100 °C in the presence of 4-anisoyl chloride. Refluxing in propan-2-ol afforded compound **(53)** in good yield. Alkylation of compound **(53)** with *tert*-butyl bromoacetate afforded ester **(55)** in poor yield. The ester was purified by heating with activated charcoal and isolated as a colourless solid. The low isolated yield of compound **(54)** was attributed to the difficulties in manipulation of compound **(53)** due to its poor solubility. The reaction was repeated by heating to 70 °C for 4 hours. However this yielded esters **(56)** and **(57)**. Removal of the *tert*-butyl ester of **(54)** with TFA afforded acid **(55)**.



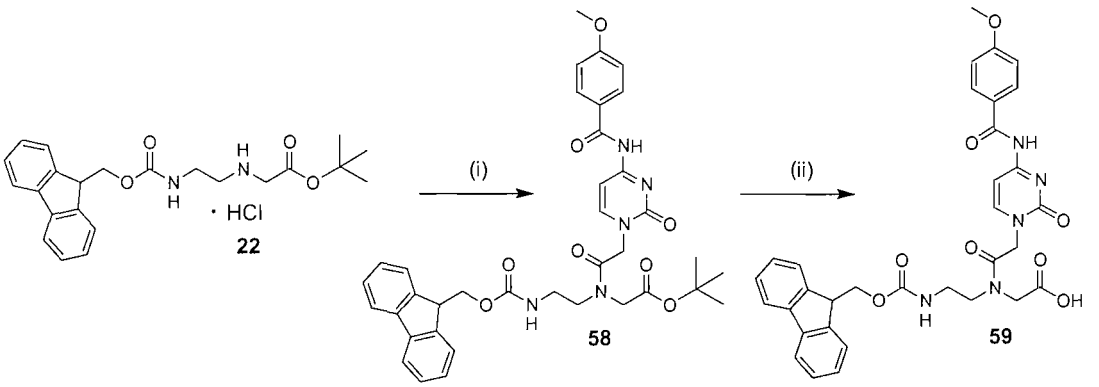
Reagents and conditions i) 4-Anisoyl chloride (1.2 eq), pyridine, 100 °C, 4 h, 84 %. ii) *tert*-Butyl bromoacetate (1.0 eq), DMF, rt, 48 h, 12 %. iii) CH_2Cl_2:triethylsilane:TFA (7:1:3 v/v/v), rt, 7 h, 99 %. iv) *tert*-Butyl bromoacetate (1.1 eq), DMF, 60 °C, 4 h, **(56)** 24 %, **(57)** 26 %.

Scheme 4.12. Synthesis of the protected adenine acetic acid (55).

4.5.2. Assembly of PNA Monomers

Having prepared the base acetic acids **(51)** and **(55)** and the Fmoc-protected PNA backbone **(22)** the synthesis of the orthogonally protected PNA monomers was attempted according to the method of Thompson *et al.* (**Scheme 4.13**).[160] Under the conditions of the coupling reaction the secondary amine on the PNA backbone **(22)** slowly removes the Fmoc group resulting in a poor yield of the desired product **(58)**.

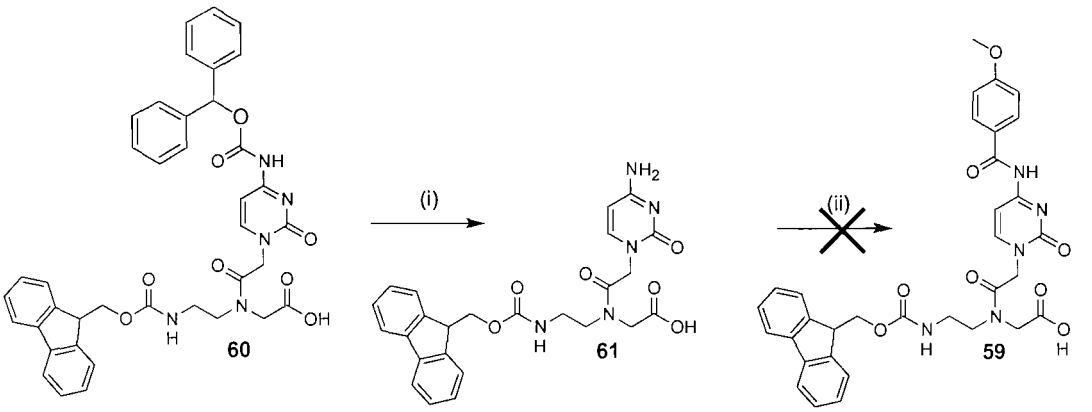
After a number of attempts this method was deemed unsuitable for production of the desired quantities of the PNA monomers, so an alternative method was investigated.



Reagents and conditions i) Acid **(51)** (1.1 eq), EDC (2.5 eq), DIPEA (1.0 eq), DMF, rt, 3 h, 13 %. ii) TFA:CH_2Cl_2 (1:1), rt, 1 h, 99 %.

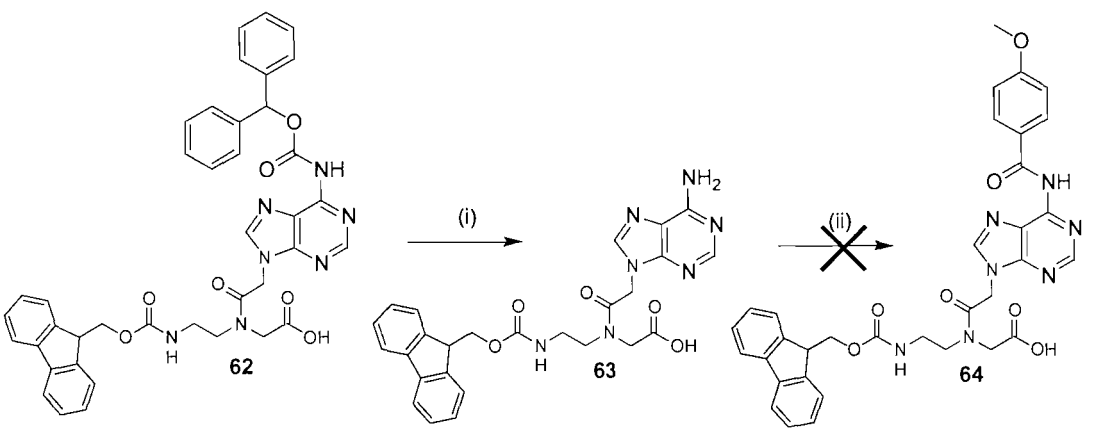
Scheme 4.13. Synthesis of the Fmoc-protected cytosine PNA monomer (59).

The PNA monomers **(60)** and **(62)** with the acid labile Bhoc protecting group on the exocyclic amines were purchased from Applied Biosystems. The Bhoc group was removed in high yields by TFA in the presence of *m*-cresol (**Schemes 4.14** and **4.15** respectively for cytosine and adenine monomers respectively). Attempts to install the anisoyl protecting group on the exocyclic amine in either pyridine or *N,N*-dimethylformamide (DMF) in the presence of DIPEA were unsuccessful. Due to time restrictions no further attempts were made to prepare the Fmoc-protected PNA monomers.



Reagents and conditions i) TFA:*m*-cresol (4:1), rt, 2 h, 95 %. ii) 4-Anisoyl chloride (1.2 eq), DIPEA, DMF or 4-Anisoyl chloride (1.2 eq), pyridine.

Scheme 4.14. Synthesis of the Fmoc-protected cytosine monomer (59).



Reagents and conditions i) TFA:*m*-cresol (4:1), rt, 2 h, 96 %. ii) 4-Anisoyl chloride (1.2 eq), DIPEA, DMF or 4-Anisoyl chloride (1.2 eq), pyridine.

Scheme 4.15. Synthesis of the Fmoc-protected adenine PNA monomer (64).

5. Conclusions

The aim of this project was to develop a mass spectrometry based hybridisation assay for SNP analysis and mutation detection. After a thorough review of literature on current mass spectrometry based protocols and methods of genotyping, a procedure which utilises PNA in a hybridisation assay followed by detection by mass spectrometry has been developed.

The assay employs PNA, as opposed to DNA due to its superior properties. Firstly, PNA is significantly easier to analyse and detect by mass spectrometry over native DNA as a consequence of the uncharged backbone. It does not readily form cation adducts which hamper sensitivity and resolution. PNA is peptide based and is synthesised by solid phase procedures. This allows peptide like modifications such as the introduction of modified monomers and amino acids and quaternary amines to provide increased sensitivity for detection by mass spectrometry.

Four different types of PNA probe have been prepared from one common scaffold.

- Unmodified PNA probe.
- Charge tagged PNA probe.
- PNA probe modified with a charge tagged mass marker.
- PNA probe modified with a charge tagged photo-cleavable mass marker.

The mass markers were composed of inexpensive, commercially available amino acids and isotopically labelled POA monomers. The POA monomers were readily synthesised from the Fmoc protected PNA backbone and were conjugated to the amino terminus of the PNA probe during synthesis on the solid phase. The function of the mass marker is to enhance the mass difference between each probe, and served as a unique molecular mass identifier in the case of the PNA probe modified with a photo-cleavable mass marker.

A basic assay protocol has been developed in which all four types of probe can be used without alterations to the assay conditions. The assay has been developed with the detection of multiple mutations (multiplexing) in mind in is amenable to liquid handling systems and high throughput applications.

Cystic fibrosis mutation analysis was chosen to evaluate the PNA hybridisation assay. In particular studies have focused on mutations on the W1282X locus. The results have highlighted a number of limitations to the assay. Firstly, a problem encountered with all hybridisation assays, stable mismatches present difficulties. Despite the superior selectivity of PNA for mismatches over conventional DNA probes, full discrimination was not always achieved due to the small differences in T_m between the fully complementary duplexes and those containing a single stable mismatch. In addition, the formation of triplexes was also observed between PNAs predominately composed of pyrimidines against homopurine DNA targets. Whist the formation of triplexes in the assay does not directly impose a problem it is not ideal and sequences which are likely to form triplexes should be avoided in future studies.

The sensitivity of the hybridisation assay was evaluated with biotinylated single stranded DNA targets. Under the generic assay conditions, excellent sensitivity was achieved with the preformed positive charge or charge tag giving an additional ten-fold increase in the sensitivity of the assay over the unmodified probe. Results showed it was necessary to derivatise the mass marker with a quaternary ammonium moiety in order to detect the mass marker cleaved from the PNA probe during laser ablation. This was readily incorporated into the synthesis of the probes.

Finally, the assay was evaluated with wild type PCR amplified DNA. All four different types of probe were used in the assay to simulate a **G** to **A** mutation and a **G** to **T** transversion mutation. The probes were successfully detected with good mass accuracy to give an unambiguous identification of the genotype of the sample by MALDI-TOF-MS.

In conclusion, a hybridisation assay has been developed which employs mass spectrometry detection. Four different types of probe have been synthesised and each type of probe has been used in the assay. All four types of probe have been used to detect mutations on the W1282X locus. Results showed that mass markers have successfully been used to enhance the mass difference between variants giving unambiguous results and rendering the assay amenable to lower specification instruments. The use of a photo-cleavable mass marker, to serve as a unique molecular mass identifier, brings the mass window of interest into the low mass region. The ability to use monoisotopic resolution, permits the use of isotope labels and would also allow the use of isotope pattern recognition software. The addition of the preformed positive charge ensures the sensitivity of detecting the cleaved mass marker is comparable to the intact probes.

6. Experimental

6.1. General

Reagents were purchased from Aldrich, Applied Biosystems, Avocado, Cruachem, Fisher, Fluka, Lancaster, Link Technologies, Nova Biochem or Rathburns and used without purification with the exception of CH_2Cl_2 and DIPEA, which were purified by distillation over calcium hydride.

6.2. Experimental for Chapter 2

6.2.1. Synthesis of PNA Probes

The PNA probes in **Table 6.1** were synthesised on a PerSeptive Biosystems Expedite nucleic acid/PNA synthesiser modified for PNA chemistry. The synthesis was performed on a 2 μM scale by the Fmoc strategy using PNA monomers and reagents purchased from Applied Biosystems. The Fmoc-protected amino acids and photo-labile linker were purchased from Nova Biochem. The crude PNA product was removed from the universal PNA column by treatment with 4:1 TFA:*m*-cresol. Following precipitation with cold ether the precipitate was washed a further 5 times with ether.

The crude samples were dried and re-dissolved in 1.5 mL water (0.1 % TFA) and 0.75 mL aliquots were purified by RP-HPLC. Purification was carried out on a Gilson instrument using an Brownlee Prep 10 Octyl column, 10 mm x 250 mm, particle size 20 μm. The system was controlled by Gilson 7.12 software and the following protocol was used: Run time 25 minutes, flow rate 4 mL min^{-1}, binary system, gradient: time in minutes (% buffer B); 0 (0); 18 (40); 21(95); 22 (95); 23 (0); 24.9 (0); 25 (0). Elution buffer A water (0.1 % TFA), buffer B acetonitrile (0.08 % TFA). Elution of PNAs was monitored by UV absorption at 300 nm. The optical densities of the purified PNAs were recorded at 260 nm on a Lambda 15 UV/Vis spectrophotomer and were used to calculate their concentration. The molecular mass of the PNA was verified by MALDI-TOF mass spectrometer on a Waters TofSpec2E operating in linear mode. 1 μL of matrix solution and 5 - 10 Dowex 50WX8-200 ion exchange beads were added to 1 μL of the dried sample. The matrix was a saturated solution of α-cyano-4-hydroxycinnamic acid in acetonitrile:water 50:50 (0.1 % TFA). Positive ion data was recorded using delayed extraction and an initial accelerating voltage of 20 kV. The coarse laser energy was set to 20% with fine adjustment being varied for each sample (typically 40 – 70 %). Approximately 90 - 110 laser shots were accumulated to produce a spectrum. The instrument was externally calibrated using a mixture of PNAs or commercially available peptides and small molecules (terfenadine, bradykinin, angiotensin I and II, oxybutynin chloride and reserpine), which were analysed under identical conditions.

PNA I.D.	Sequence	Calculated Average Mass (Da) $(M+H)^+/(M)^+$	O.D. mL^{-1}
PNA(14**C**)	CTTTCCT**C**CACTGT	3690.5740	37.0
PNA(14**A**)	CTTTCCT**A**CACTGT	3714.5991	34.2
PNA-M(14**C**)	Br-Ala-Lys-CTTTCCT**C**CACTGT	4186.9785	27.2
PNA-M(14**A**)	Cl-Phe-Lys-CTTTCCT**A**CACTGT	4242.6501	25.4
PNA-L-M(14**C**)	Br-Ala-Lys-hv-CTTTCCT**C**CACTGT	4467.2590	28.8
PNA-L-M(14**A**)	Cl-Phe-Lys-hv-CTTTCCT**A**CACTGT	4522.9306	12.8
PNA^+(14**C**)	$((CH_3)_3{}^+N(CH_2)_5CO)$CTTTCCT**C**CACTGT	3843.8142	ND
PNA^+(14**A**)	$((CH_3)_3{}^+N(CH_2)_5CO)$CTTTCCT**A**CACTGT	3869.8393	ND
PNA(12**C**)	CTTTCCT**C**CACT	3133.0477	33.0
PNA(12**T**)	CTTTCCT**T**CACT	3148.0593	29.2
PNA^+(12**C**)	$((CH_3)_3{}^+N(CH_2)_5CO)$CTTTCCT**C**CACT	3288.2897	ND
PNA^+(12**T**)	$((CH_3)_3{}^+N(CH_2)_5CO)$CTTTCCT**T**CACT	3303.2995	ND
PNA-M^+(12**C**)	(CH_3CO)Br-Ala-$((CH_3)_3{}^+N(CH_2)_5CO)$Lys-CTTTCCT**C**CACT	3826.7296	20.0
PNA-M^+(12**T**)	(CH_3CO)Cl-Phe-$((CH_3)_3{}^+N(CH_2)_5CO)$Lys-CTT**T**CCT**T**CACT	3873.3877	2.0
PNA-L-M^+(12**C**)	(CH_3CO)Br-Ala-$((CH_3)_3{}^+N(CH_2)_5CO)$Lys-hv-CTTTCCT**C**CACT	4107.0102	9.2
PNA-L-M^+(12**T**)	(CH_3CO)Cl-Phe-$((CH_3)_3{}^+N(CH_2)_5CO)$Lys-hv-CTTTCCT**T**CACT	4125.6346	2.7
PNA(12**A**)	CTTTCCT**A**CACT	3157.6728	38.0
PNA^+(12**A**)	$((CH_3)_3{}^+N(CH_2)_5CO)$CTTTCCT**A**CACT	3312.3219	ND
PNA-M^+(12**A**)	(CH_3CO)Cl-Phe-$((CH_3)_3{}^+N(CH_2)_5CO)$Lys-CTTTCCT**A**CACT	3882.4012	12.8
PNA-L-M^+(12**A**)	(CH_3CO)Cl-Phe-$((CH_3)_3{}^+N(CH_2)_5CO)$Lys-hv-CTTTCCT**A**CACT	4162.6817	11.9

Table 6.1. List of all PNAs synthesised.

6.2.2. Synthesis of Oligonucleotides

All oligonucleotides were synthesised on an Applied Biosystems 394 automated DNA/RNA synthesiser using the standard 0.2 μmole phosphoramidite cycle of acid-catalysed detritylation, coupling, capping and iodine oxidation. Stepwise coupling efficiencies and overall yields were determined by the automated trityl cation conductivity monitoring facility and in all cases were >98 %. All β-cyanoethyl phosphoramidite monomers were dissolved in anhydrous acetonitrile to a concentration of 0.1 M immediately prior to use. Biotin was added with a biotin TEG monomer purchased from Glen Research Inc. The biotinylated oligonucleotides were purified by HPLC prior to removal of the 5'-trityl-protecting group with 80% acetic acid/20% water at room temperature for 30 minutes. Standard DNA phosphoramidites, solid supports and additional reagents were purchased from Link Technologies or Applied Biosystems. Purification was carried out by RP-HPLC on a Gilson system using an ABI Aquapore column (C8), 8 mm x 250 mm, pore size 300 Å. The system was controlled by Gilson 7.12 software and the following protocol was used: Run time 30 minutes, flow rate 4 mL min^{-1}, binary system, gradient: time in minutes (% buffer B);0 (0); 3 (0); 5 20); 21 (100); 25 (100); 27 (0); 30 (0). Elution buffer A, 0.1 M ammonium acetate, pH 7.0, buffer B, 0.1 M ammonium acetate (35% acetonitrile) pH 7.0. Elution of oligonucleotides was monitored by UV absorption at 295 nm. After purification, oligonucleotides were desalted using disposable NAP 10 Sephadex columns (Pharmacia) and their optical densities recorded on a Lambda 15 UV/Vis spectrophotomer. The molecular mass of the oligonucleotides were verified by MALDI-TOF mass spectrometer on a Waters TofSpec2E operating in linear mode. 1 μL of matrix solution and 5 - 10 Dowex 50WX8-200 ion exchange beads were added to 1 μL of the dried sample.[57] The matrix was a saturated solution of 4:1 HPA:PA in 1:1 acetonitrile:water. Positive ion data was recorded using delayed extraction and an initial accelerating voltage of 20 kV.

The coarse laser energy was set to 50% with fine adjustment being varied for each sample (typically 40 – 70 %). Approximately 90 - 110 laser shots were accumulated to produce a spectrum. The instrument was externally calibrated using a mixture of synthetic oligonucleotides, which were analysed under identical conditions.

DNA I.D.	Sequence	Calculated Average Mass (Da) $(M+H)^+$	O.D. mL^{-1}
Primer	GGCTAAGTCCTTTTGCTCAC	6058.9830	7.7
Primer	ATGGTGTGTCTTGGGATTCA-Biotin	6762.6737	3.7
	ACAGTG**T**AGGAAAG	4360.9167	14.4
	ACAGTG**A**AGGAAAG	4361.9246	43.4
	ACAGTG**G**AGGAAAG	4385.9296	8.5
	ACAGTG**C**AGGAAAG	4346.9130	8.9
	ACAG**T**GTAGGAAAG-Biotin	4939.6018	5.9
	ACAGTG**A**AGGAAAG-Biotin	4930.5303	24.0
	ACAGTG**G**AGGAAAG-Biotin	4965.6225	4.8

Table 6.2. List of all oligonucleotides synthesised.

6.2.3. UV-Melting

The UV absorbance *vs.* temperature profiles were obtained simultaneously at 260 nm using an intra-cuvette temperature probe on a ChemCary 400 UV Visible spectrometer in Hellma® SUPRASIL synthetic quartz, 10 mm path length cuvettes. The PNA probe and oligonucleotide target were each at a concentration of 1.0 μM in 10 mM Tris, 1 M NaCl, pH 7 in a volume of 1.5 mL.

The samples were filtered with Kinesis regenerated cellulose 13 mm, 0.45 μM syringe filters and rapidly heated from 15 °C to 80 °C at 10 °C min^{-1} then cooled to 15 °C at 0.5 °C min^{-1} prior to analysis. The UV melting curves were recorded for three consecutive heat and cool cycles. The temperature was increased in increments of 0.1 °C min^{-1} and the T_m values were determined from the maximum of the first order derivative of the average of the three cycles.

6.2.4. PNA Hybridisation Assay

The purified PCR product was added to 30 μg of Dynal Biotech M-280 streptavidin beads and was allowed to bind at room temperature for 30 minutes in 20 μL of binding buffer (10 mM Tris, 1 M NaCl, pH 7). The beads were isolated and washed with 0.1 M NaOH and re-suspended in 200 μL of 0.1 M NaOH. After 5 minutes the beads were isolated and washed twice with 200 μL of binding buffer and twice with 200 μL of hybridisation buffer (10 mM Tris, 0.1 % BSA, 1 M NaCl, pH 7). For hybridisation 30 pmol of PNA was added to 20 μL of hybridisation buffer and the mixture was incubated at room temperature for 15 minutes. The mixture was heated to 80 °C for 1 hour and allowed to cool to room temperature over 1 hour.

The beads were isolated and washed twice with 200 μL of hybridisation buffer and twice with binding buffer then re-suspended in 200 μL of binding buffer and incubated at 45 °C for 20 minutes. They were then washed twice with 200 μL of binding buffer, twice with 200 μL of wash buffer (10 mM Tris, 0.1 % SDS, pH 7) and twice with 200 μL of water then re-suspended in 1 μL of matrix (α-cyano-4-hydroxycinnamic acid in acetonitrile:water 50:50 (0.1 % TFA) and loaded directly on the MALDI plate with Dowex 50WX8-200 ion exchange beads. Spectra were collected using a Waters TofSpec2E as described previously.

6.2.5. PCR Amplification of Human DNA

The CFTR W1282X region was amplified using one 5' biotinylated primer (**ATGGTGTGTCTTGGGATTCA**) and one unmodified primer (**GGCTAAGTCCTTTTGCTCAC**). PCR amplifications were performed using a Progene thermal cycler (Techne). PCR reagents and buffers were obtained from Promega, except human genomic DNA, which was purchased from Novagen. PCR reactions in a final volume of 50 μL contained each primer (0.5 μM), dNTPs (200 μM each of dATP, dCTP, dTTP and dGTP), $MgCl_2$ (3 mM), Taq Polymerase (2.5U) and 1x PCR Buffer. Human DNA template (5 ng μL^{-1}) or water for no template controls, was added to each reaction. Each tube was subjected to an initial activation (95°C, 5 minutes), and then subjected to 35 cycles of annealing (65 °C, 30 seconds), extension (65 °C, 30 seconds) and denaturing (74 °C, 30 seconds), followed by a final extension phase (72 °C, 10 minutes). After thermal cycling, 5 μL of each sample was electrophoresed on a 2 % agarose gel stained with ethidium bromide, and photographed under a UV-transilluminator. Samples were then individually purified using a Qiagen PCR purification kit and concentrated to 10 μL.

6.2.6. Derivatisation of PNA probes on MALDI Target

1 μL (3 mg, 3.6 μmol) of C_5Q in 1 M trimethylammonium bicarbonate (0.5 mL) at pH 8.5 and 5 °C was added to the PNA/DNA bead complex on the MALDI platen. The mixture was left for 10 minutes at 5 °C and samples were then dried *in vacuo*.

6.3. Experimental for Chapter 3

6.3.1. Synthesis of PNA Probes

PNA probes were prepared as described previously, however, prior to deprotection from the solid phase 1 mL of PNA capping reagent (5 % acetic anhydride, 6 % 2-6 lutidine and DMF) was added to the PNA on the solid phase. The mixture was left for 10 minutes at room temperature and was washed with CH_2Cl_2 (25 mL) and dried by flushing with argon. The PNA was then removed from the solid phase as described previously. C_5Q (0.25 % w/v) in 1 M trimethylammonium bicarbonate pH 8.5 at 5 °C (0.75 mL) was added to the dried PNA. The mixture left for 2 hours at 5 °C, then the samples were lyphilised to dryness. The crude samples were re-dissolved in 1.5 mL water (0.1 % TFA) and 0.75 mL aliquots were purified by RP-HPLC as described previously.

6.3.2. PNA Hybridisation Assay

The purified PCR product was added to 30 μg of Dynal Biotech M-280 streptavidin beads and was allowed to bind at room temperature for 30 minutes in 20 μL of binding buffer (10 mM Tris, 1 M NaCl, pH 7). The beads were isolated and washed with 0.1 M NaOH and re-suspended in 200 μL of 0.1 M NaOH. After 5 minutes the beads were isolated and washed twice with 200 μL of binding buffer and twice with 200 μL of hybridisation buffer (10 mM Tris, 0.1 % BSA, 1 M NaCl, pH 7). For hybridisation 30 pmol of PNA was added to 20 μL of hybridisation buffer and the mixture was incubated at room temperature for 15 minutes. The mixture was heated to 80 °C for 1 hour and allowed to cool to room temperature over 1 hour.

The beads were isolated and washed twice with 200 μL of hybridisation buffer and twice with binding buffer then re-suspended in 200 μL of binding buffer and incubated at 55 °C (PNA probes with **C** and **T** at variant site) or 30 °C (PNA probes with **C** and **A** at variant site) for 20 minutes for PNA probes designed to detect **G** to **A** mutation and 30 °C for 20 minutes designed to detect **G** to **T** mutation. They were then washed twice with 200 μL of binding buffer, twice with 200 μL of wash buffer (10 mM Tris, 0.1 % SDS, pH 7) and twice with 200 μL of water then re-suspended in 1 μL of matrix (α-cyano-4-hydroxycinnamic acid in acetonitrile:water 50:50 (0.1 % TFA) and loaded directly on the MALDI platen with Dowex 50WX8-200 ion exchange beads. Spectra were collected using a Waters TofSpec2E as described previously.

6.3.3. Circular Dichroism Spectroscopy

The oligonucleotide target was at a concentration of 24 μM in 10 mM Tris, 1 M NaCl, pH 7 in a volume of 250 μL. The samples were filtered into the cuvettes with Kinesis regenerated cellulose 13 mm, 0.45 μM syringe filters and aliquots of PNA were added. CD experiments were recorded on a JASCO J-710 spectropolarimeter at 25 °C in a 1 mm diameter cuvette from 210-340 nm with a resolution of 0.2 nm, at a scan speed of 100 nm min^{-1}, response of 4 seconds and a bandwidth of 1.0 nm. Each spectrum was accumulated 5 times, smoothed and the spectrum of the buffer was subtracted.

6.3.4. Bruker Ultraflex MALDI-TOF

Data was recorded using a Bruker Ultraflex MALDI-TOF mass spectrometer operating in reflectron mode.

Positive ion data were recorded using delayed extraction and an initial accelerating voltage of 20 kV in α-cyano-4-hydroxycinnamic acid in acetonitrile:water 50:50 (0.1 % TFA) using Dowex 50WX8-200 ion exchange beads. The instrument was calibrated using a mixture commercially available peptides or PNAs.

6.4. Experimental for Chapter 4

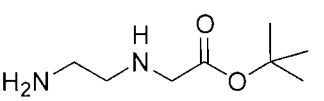
6.4.1. General

All chemical reactions were carried out under argon using oven-dried glassware. Column chromatography was carried out under pressure using Fisher scientific DAVISIL 60A (35-70 micron) silica. Compounds were visualised by irradiation at 254 nm or by staining with phosphomolybdic acid:C_2H_5OH (1:10). Thin layer chromatography was performed using Merck Kieselgel 60 F24 (0.22 mm thickness, aluminium backed). ^1H NMR spectra were obtained at 300MHz on a Bruker AC 300 spectrometer or at 400 MHz on a Bruker DPX400 spectrometer and ^{13}C NMR spectra were measured at 75 or 100 MHz on the same spectrometer. Chemical shifts are given in ppm relative to the residual solvent peak, and *J* values are given in Hz. Low-resolution mass spectra were recorded using electrospray ionisation on a Waters ZMD quadrupole mass spectrometer in methanol or water. High-resolution mass spectra were recorded on a Bruker APEX III FT-ICR mass spectrometer in methanol or water using an Apollo electrospray ionisation source. Infrared spectra were recorded on a BIORAD FT-IR using a Golden Gate adapter and BIORAD WIN-IR software or a Satellite FT-IR using a Golden Gate adapter and WIN FIRST-lite software. Absorptions are described as strong (s), medium (m), broad (br) or weak (w).

tert-Butyl 2-[(2-aminoethyl)amino]acetate (13), _tert_-butyl 2-[(2-[2-(_tert_-butoxy)-2-oxoethyl]aminoethyl)amino]acetate (14), _tert_-butyl 2-[[2-(_tert_-butoxy)-2-oxoethyl](2-[2-(_tert_-butoxy)-2-oxoethyl]aminoethyl)amino]acetate (15)

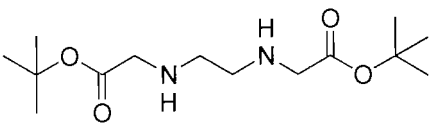
Following the procedure described by Thompson *et al.*[160] *tert*-Butyl bromoacetate (8 mL, 56 mmol) in CH_2Cl_2 (200 mL) was added over 5 hours to a rapidly stirring solution of ethylene diamine **(12)** (30 mL, 449 mmol) in CH_2Cl_2 (40 mL) at 0 °C. The mixture was allowed to warm to room temperature and was stirred overnight. The reaction mixture was washed with water (3 x 500 mL) and the aqueous layers back extracted with CH_2Cl_2 (250 mL). The organic layers were combined, dried over Na_2SO_4, filtered and the solvent was removed *in vacuo.* The resulting oil was triturated with Et_2O yielding amine **(12)** as a white solid (0.41 g, 42 %). Compound **(14)** (0.96 g, 6 %) and compound **(15)** (0.45 g, 2 %) were isolated as colourless oils from the Et_2O by column chromatography (CH_2Cl_2 to 6:4 CH_2Cl_2:CH_3OH).

Data for _tert_-butyl 2-[(2-aminoethyl)amino]acetate (13)



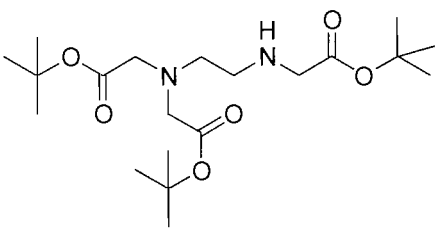
Data: R_f = 0.62 (10:1 CH_2Cl_2:CH_3OH) UV and ninhydrin; υ_{max} (Film, cm^{-1}) 2977 (m), 2931 (w), 1731 (s), 1561 (w), 1459 (m), 1420 (w), 1392 (m), 1367 (m), 1292 (w), 1227 (m), 1155 (s); δ_H (300 MHz, $CDCl_3$) 3.35 (2H, s, $NHCH_2CO_2C(CH_3)_3$), 2.75 (2H, t, *J*=5.9 Hz, $NH_2CH_2CH_2NH$), 2.65 (2H, t, *J*=5.9 Hz, $H_2NCH_2CH_2NH$), 1.50 (9H, s, $C(CH_3)_3$); δ_C (75 MHz, $CDCl_3$) 172.06 (0), 81.28 (0), 52.24 (2), 51.71 (2), 41.87 (2), 28.23 (3); LRMS (ES+) *m/z* 175.0 $(M+H)^+$ (100 %).

Data for *tert*-butyl 2-[(2-[2-(*tert*-butoxy)-2-oxoethyl]aminoethyl)amino]acetate (14)



Data: R_f = 0.57 (10:1 CH_2Cl_2:CH_3OH) UV and ninhydrin; δ_H (300 MHz, $CDCl_3$) 3.30 (4H, s, NHC**H**$_2$CO$_2$C(CH$_3$)$_3$), 2.70 (4H, s, NHC**H**$_2$C**H**$_2$NH), 1.45 (18H, s, C(CH$_3$)$_3$); δ_C (75 MHz, $CDCl_3$) 171.98 (0), 81.22 (0), 51.74 (2), 49.06 (2), 28.26 (3); LRMS (ES+) *m/z* 289.2 $(M+H)^+$ (50 %), 311.2 $(M+Na)^+$ (20 %), 233.2 $(M+H-C_4H_8)^+$ (100 %).

Data for *tert*-butyl 2-[[2-(*tert*-butoxy)-2-oxoethyl](2-[2-(*tert*-butoxy)-2-oxoethyl]aminoethyl)amino]acetate (15)

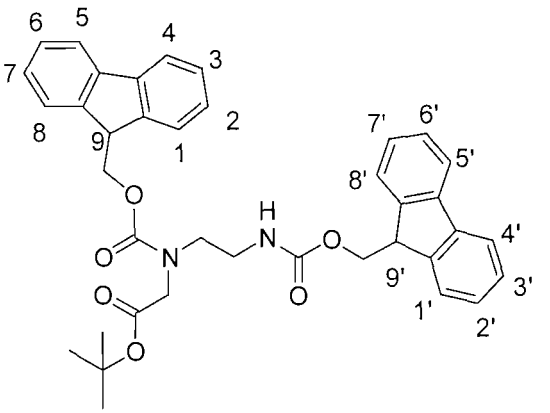


Data: R_f = 0.70 (10:1 CH_2Cl_2:CH_3OH) UV and ninhydrin; υ_{max} (film, cm^{-1}) 2977. (w), 2932 (w), 2360 (w), 2336 (w), 1731 (s), 1476 (w), 1457 (w), 1392 (w), 1367 (m), 1238 (w), 1249 (w), 1219 (w), 1147 (s); δ_H (300 MHz, $CDCl_3$) 3.50 (4H, s, NC**H**$_2$CO$_2$C(CH$_3$)$_3$), 3.30 (2H, s, NHC**H**$_2$CO$_2$C(CH$_3$)$_3$), 2.85 (2H, t, *J*=6.3 Hz, NCH$_2$C**H**$_2$NH), 2.65 (2H, t, *J*=6.3 Hz, NC**H**$_2$CH$_2$NH), 2.05 (1H, br s, N**H**), 1.50 (9H, s, C(C**H**$_3$)$_3$), 1.45 (18H, s, C(C**H**$_3$)$_3$); δ_C (75 MHz, $CDCl_3$) 171.76 (0), 170.98 (0), 81.05 (0), 80.85 (0), 56.17 (2) 54.07 (2), 51.81 (2), 47.28 (2), 28.30 (3), 28.26 (3); LRMS (ES+) *m/z* 403.3 $(M+H)^+$ (100 %), 347.2 $(M+H-C_4H_8)^+$ (45 %); HRMS (ES+) *m/z* 403.2793 $(M+H)^+$ $C_{20}H_{39}N_2O_6$ requires 403.2803.

tert-Butyl 2-[[(9_H_-9-fluorenylmethoxy)carbonyl](2-[(9'_H_-9'-fluorenylmethoxy)carbonyl]aminoethyl)amino]acetate (16), 9_H_-9-fluorenylmethyl 3-[2-(_tert_-butoxy)-2-oxoethyl]-1-imidazolidinecarboxylate (17) and _tert_-butyl 2-(2-[2-(_tert_-butoxy)-2-oxoethyl][(9_H_-9-fluorenylmethoxy)carbonyl]aminoethyl)[(9_H_-9-fluorenylmethoxy)carbonyl]aminoacetate (18)

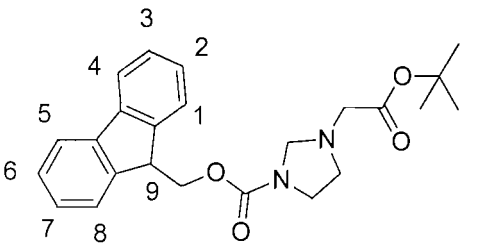
Following the procedure described by Thompson *et al.*[160] Fmoc-Su (4.54 g, 13.4 mmol) in CH_2Cl_2 (80 mL) was added over 15 minutes to a rapidly stirring solution of amine **(13)** (2.34 g, 13.4 mmol) in CH_2Cl_2 (300 ml) and DIPEA (2.28 mL, 13.4 mmol). After 45 minutes the reaction mixture was washed with 1 M HCl (3 x 200 mL) then brine (3 x 200 mL). The organic layer was dried over Na_2SO_4 and filtered. The mixture was concentrated *in vacuo* and purified by silica column chromatography (2:8 ethyl acetate:hexane to 1:1 ethyl acetate:hexane). Fractions containing products were concentrated *in vacuo* to yield compounds **(16)** (2.34 g, 28 %), **(17)** (0.68 g, 13 %) and **(18)** 0.48 g, 5 %) as colourless foams.

Data for *tert*-butyl 2-[[(9*H*-9-fluorenylmethoxy)carbonyl](2-[(9*H*'-9'-fluorenylmethoxy)carbonyl]aminoethyl)amino]acetate (16)



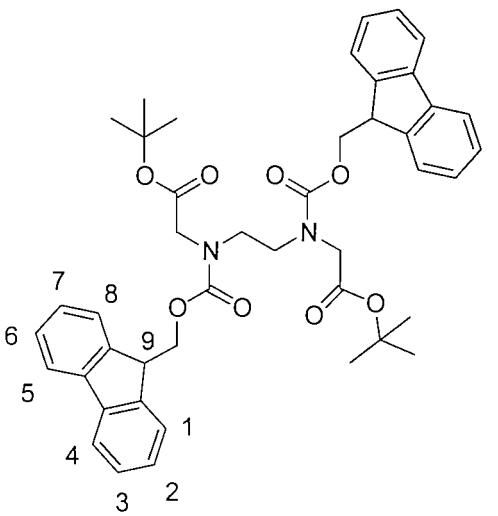
Data: R_f = 0.55 (5:1:1 ethyl acetate:CH_3OH:NH_3) UV; υ_{max} (Film, cm^{-1}) 3065 (w), 2977 (w), 2944 (w), 1705 (s), 1702 (s), 1608 (w), 1517 (m), 1477 (m), 1450 (m), 1430 (m), 1407 (w), 1394 (w), 1367 (m), 1317 (w), 1240 (s), 1217 (m), 1153 (s), 1105 (w), 1081 (w); δ_H (400 MHz, $CDCl_3$) (1:1 rotamers) 7.72-7.80 (8H, m, $\mathbf{H}^{4/5}$, $\mathbf{H}^{4'/5'}$), 7.51-7.62 (8H, m, $\mathbf{H}^{1/8}$, $\mathbf{H}^{1'/8'}$), 7.35-7.48 (8H, m, $\mathbf{H}^{3/6}$, $\mathbf{H}^{3'/6'}$), 7.18-7.30 (8H, m, $\mathbf{H}^{2/7}$, $\mathbf{H}^{2'/7'}$), 5.68 (1H, br t, N**H**), 5.21 (1H, br t, N**H**), 4.54-4.60 (4H, m, $C\mathbf{H}_2CH^9$, $C\mathbf{H}_2CH^{9'}$), 4.30-4.43 (4H, m, $C\mathbf{H}_2CH^9$, $C\mathbf{H}_2CH^{9'}$), 4.26-4.13 (4H, m, $C\mathbf{H}^9$, $C\mathbf{H}^{9'}$), 3.84 (2H, s, $NC\mathbf{H}_2CO_2C(CH_3)_3$), 3.68 (2H, s, $NC\mathbf{H}_2CO_2C(CH_3)_3$), 3.41-3.48 (2H, m, $NCH_2C\mathbf{H}_2NH$), 3.33-3.40 (2H, m, $NC\mathbf{H}_2CH_2NH$), 3.21-3.28 (2H, m, $NC\mathbf{H}_2CH_2NH$), 3.02-3.11 (2H, m, $NCH_2C\mathbf{H}_2NH$), 1.30 (18H, 2×s, $C(C\mathbf{H}_3)_3$); δ_C (100 MHz, $CDCl_3$) 144.23 (0), 143.95 (0), 141.43 (0), 140.68 (0), 127.85 (1), 127.37 (1), 127.21 (1), 125.34 (1), 125.23 (1), 124.93 (1), 124.27 (1), 120.95 (1), 82.64 (0), 68.19 (2), 67.46 (2), 66.90 (2), 66.66 (2), 51.37 (2), 51.01 (2), 49.28 (2), 47.38 (1), 39.80 (2), 39.60 (2), 28.22 (3), 28.00 (3); LRMS (ES+) *m/z* 641.4 $(M+Na)^+$ (100 %); HRMS (ES+) *m/z* 641.2636 $(M+Na)^+$; $C_{38}H_{38}N_2O_6Na$ requires 641.2622.

Data for 9*H*-9-fluorenylmethyl 3-[2-(*tert*-butoxy)-2-oxoethyl]-1-imidazolidinecarboxylate (17)



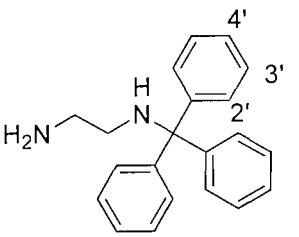
Data: R_f = 0.40 (5:1:1 ethyl acetate:CH_3OH:NH_3) UV; υ_{max} (Film, cm^{-1}) 3066 (w), 2975 (w), 2891 (w), 2812 (w), 2361 (w), 2246 (w), 1702 (s), 1648 (w), 1608 (w), 1578 (w), 1477 (m), 1450 (m), 1420 (m), 1392 (w), 1365 (m), 1345 (m), 1242 (m), 1224 (m), 1199 (m), 1147 (s), 1096 (w), 1033(w); δ_H (300 MHz, $CDCl_3$) 7.85 (2H, d, *J*=7.4 Hz, $\mathbf{H}^{4/5}$), 7.56-7.64 (2H, m, $\mathbf{H}^{1/8}$), 7.40 (2H, t, *J*=7.4 Hz, $\mathbf{H}^{3/6}$), 7.30 (2H, t, *J*=7.4 Hz, $\mathbf{H}^{2/7}$), 4.40 (2H, d, *J*=6.6 Hz, $C\mathbf{H}_2CH^9$),4.25 (1H, t, *J*=6.6 Hz, $C\mathbf{H}^9$), 3.57 (2H, s, $NC\mathbf{H}_2CO_2C(CH_3)_3$), 3.45-3.60 (2H, m, $NCH_2C\mathbf{H}_2N$), 3.33 (2H, s, $NC\mathbf{H}_2N$), 2.98-3.13 (2H, m, $NC\mathbf{H}_2CH_2N$), 1.50 (9H, s, $C(C\mathbf{H}_3)_3$); δ_C (75 MHz, $CDCl_3$) 169.23 (0), 154.02 (0), 144.11 (0), 141.46 (0), 127.87 (1), 127.21 (1), 125.23 (1), 120.14 (1), 81.86 (0), 68.10 (2), 55.45 (2), 52.26 (2), 47.39 (1), 44.24 (2), 44.14 (2), 28.25 (3); LRMS (ES+) *m/z* 431.4 $(M+Na)^+$ (100 %); LRMS (ES+ in CD_3OD) *m/z* 410.2 $(M+D)^+$ (100 %), HRMS (ES+) *m/z* 409.2122 $(M+H)^+$; $C_{24}H_{29}N_2O_4$ requires 409.2123.

Data for *tert*-butyl 2-(2-[2-(*tert*-butoxy)-2-oxoethyl][(9*H*-9-fluorenylmethoxy)carbonyl]aminoethyl)[(9*H*-9-fluorenylmethoxy)carbonyl]aminoacetate (18)



Data: R_f = 0.70 (5:1:1 ethyl acetate:CH_3OH:NH_3) UV; υ_{max} (Film, cm^{-1}) 3066 (w), 2977 (w), 2251 (w), 1741 (s), 1701 (s), 1608 (w), 1475(m), 1450 (m), 1431 (m), 1407 (w), 1393 (w), 1366 (m), 1315 (m), 1221 (s), 1150 (s), 1113 (s), 1033(w); δ_H (300 MHz, $CDCl_3$) (2:1 rotamers) 7.69-7.75 (12H, m, $\mathbf{H}^{4/5}$), 7.48-7.56 (12H, m, $\mathbf{H}^{1/8}$), 7.25-7.32 (12H, m, $\mathbf{H}^{3/6}$), 7.16-7.26 (12H, m, $\mathbf{H}^{2/7}$), 4.56 (8H, d, *J*=6.7, $C\mathbf{H}_2CH^9$), 4.43 (4H, d, *J*=6.7, $C\mathbf{H}_2CH^9$), 4.21-4.00 (6H, m, $C\mathbf{H}^9$), 3.87 (8H, s, $NC\mathbf{H}_2CO_2C(CH_3)_3$), 3.56 (4H, s, $NC\mathbf{H}_2CO_2C(CH_3)_3$), 3.22-3.33 (8H, m, NCH_2CH_2N), 3.16-3.20 (4H, m, $NC\mathbf{H}_2C\mathbf{H}_2N$), 1.36 (36H, s, $C(C\mathbf{H}_3)_3$), 1.36 (18H, s, $C(C\mathbf{H}_3)_3$); δ_C (75 MHz, $CDCl_3$) 168.96 (0), 156.35 (0), 143.98 (0), 141.39 (0), 127.90 (1), 127.37 (1), 127.24 (1), 125.25 (1), 125.15 (1), 124.94 (1), 124.87 (1), 120.12 (1), 82.14 (0), 68.10 (2), 67.34 (2), 67.22 (2), 65.93 (2), 51.03 (1), 47.93 (2), 47.36 (2), 46.86 (2), 46.64 (2), 28.24 (3), 28.00 (3); LRMS (ES+) *m/z* 755.5 $(M+Na)^+$ (100 %); HRMS (ES+) *m/z* 771.3070 $(M+K)^+$; $C_{44}H_{48}N_2O_8K$ requires 771.3043.

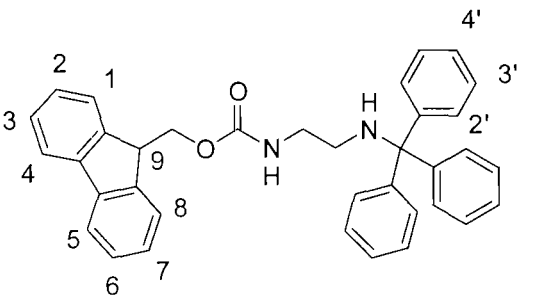
***N*-1-Trityl-1,2-ethanediamine (19)**



Trityl chloride (15.00 g, 53.81 mmol) in anhydrous pyridine (250 mL) was added drop wise to a rapidly stirred solution of ethylene diamine **(12)** (18.51 mL, 269.03 mmol) in anhydrous pyridine (50 mL). The reaction was stirred rapidly for 4 hours under an atmosphere of argon. The mixture was evaporated to dryness (azeotroping with toluene) and the residue was dissolved in CH_2Cl_2 (50 mL) and washed with water (2 x 200 mL). The organic layer was dried over Na_2SO_4 and concentrated *in vacuo* yielding amine **(19)** as a colourless oil (15.90 g, 97 %).

Data: R_f = 0.67 (5:1:1 ethyl acetate:CH_3OH:NH_3) UV and ninhydrin; υ_{max} (Film, cm^{-1}) 3080 (w), 3055 (w), 3029 (w), 1595 (m), 1489 (s), 1448 (s), 1316 (w), 1210 (w), 1183 (w), 1156 (w), 1108 (w), 1082 (w), 1032 (w), 1001 (w); δ_H (300 MHz, d_6-DMSO) 7.45 (6H, d, *J*=7.3 Hz, $\mathbf{H}^{2'}$), 7.35 (6H, t, *J*=7.3 Hz, $\mathbf{H}^{3'}$), 7.15 (3H, d, *J*=7.3 Hz, $\mathbf{H}^{4'}$), 3.05 (1H, br s, N**H**), 2.60 (2H, t, *J*=6.6 Hz, $C\mathbf{H}_2NH$), 2.00 (2H, t, *J*=6.6 Hz, $C\mathbf{H}_2NH_2$), 1.65 (2H, br s, $N\mathbf{H}_2$); δ_C (75 MHz, d_6-DMSO) 146.40 (0), 128.52 (1), 127.74 (1), 126.06 (1), 70.26 (0), 46.93 (2), 2.25 (2); LRMS (ES+) *m/z* 303.3 $(M+H)^+$ (10 %), 243.2 $(C_{19}H_{16})^+$ (100 %); HRMS (ES+) *m/z* 303.1854 $(M+H)^+$; $C_{21}H_{23}N_2$ requires 303.1856.

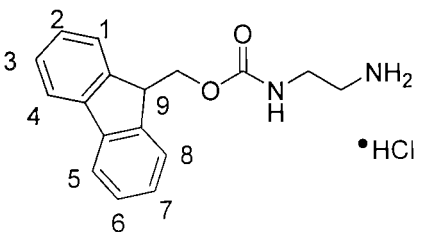
9*H*-9-Fluorenylmethyl *N*-[2-(tritylamino)ethyl]carbamate (20)



Fmoc-Su (30.15 g 89.38 mmol) in CH_2Cl_2 (100 mL) was added over 1 hour to a rapidly stirring solution of amine **(19)** (27.03 g, 89.38 mmol) and DIPEA (8.6 mL, 89.38 mmol) in CH_2Cl_2 (500 mL). The mixture was stirred for 1 hour at room temperature. The reaction mixture was washed with water (3 x 200 mL) then brine (3 x 200 mL). The organic layer was dried over Na_2SO_4, filtered concentrated *in vacuo* yielding amide **(20)** as a colourless solid (32.90 g 70 %).

Data: R_f = 0.81 (4:1 Et_2O:ethyl acetate) UV and phosphomolybdic acid; υ_{max} (Film, cm^{-1}) 3417 (w), 3058 (w), 3020 (w), 2947 (w), 2853 (w), 1708 (s), 1595 (w), 1490 (m), 1465 (w), 1448 (s), 1409 (w), 1320 (w), 1249 (s), 1185 (w), 1143 (w), 1165 (s), 1109 (w), 1082 (w), 1032 (w), 1022 (w); δ_H (300 MHz, $CDCl_3$) 7.60 (2H, d, *J*=7.4 Hz, $\mathbf{H}^{4/5}$), 7.05 (2H, d, *J*=7.4 Hz, $\mathbf{H}^{1/8}$), 7.50 (6H, d, *J*=7.3 Hz, $\mathbf{H}^{2'}$), 7.40 (2H, t, *J*=7.4 Hz, $\mathbf{H}^{3/6}$), 7.35 (2H, t, *J*=7.4 Hz, $\mathbf{H}^{2/7}$), 7.30 (6H, t, *J*=7.3 Hz, $\mathbf{H}^{3'}$), 7.30 (3H, d, *J*=7.3 Hz, $\mathbf{H}^{4'}$), 5.10 (1H, br s N**H**CO), 4.45 (2H, d, *J*=6.6 Hz, C$\mathbf{H}_2$CH^9), 4.35 (1H, t, *J*=6.6 Hz, C$\mathbf{H}^9$), 3.35 (2H, br s, FmocNHC$\mathbf{H}_2$$CH_2$NH), 2.30 (2H, br s, NHCH$_2$C$\mathbf{H}_2$NHTrt), 1.80 (1H, br s, N**H**Trt); δ_C (75 MHz, $CDCl_3$) 156.72 (0), 145.89 (0), 144.14 (0), 141.47 (0), 128.70 (1), 128.09 (1), 127.86 (1), 127.22 (1), 126.60 (1), 125.21 (1), 120.15 (1), 71.86 (0), 66.88 (2), 47.43 (1), 43.85 (2), 42.63 (2); LRMS (ES+) *m/z* 243.2 $(C_{19}H_{16})^+$ (100 %), 525.3 $(M+H)^+$ (20 %), 547.3 $(M+Na)^+$ (5 %); HRMS (ES+) *m/z* 525.2545 $(M+H)^+$; $C_{36}H_{33}N_2O_2$ requires 525.2536.

9*H*-9-Fluorenylmethyl *N*-(2-aminoethyl)carbamate hydrochloride (21)



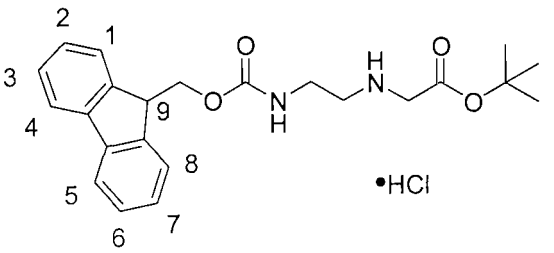
TFA (20 mL) was added drop wise to a rapidly stirring solution of amide **(20)** (4.21 g, 8.02 mmol) and triethylsilane (10 mL) in CH_2Cl_2 (50 mL) at room temperature. The mixture was stirred for 30 minutes under an atmosphere of argon. The mixture was concentrated *in vacuo* yielding the hydrochloride salt of amine **(21)** as a colourless solid, which was triturated with Et_2O and dried in air (2.26 g, 100 %).

Data: R_f = 0.50 (5:1:1 ethyl acetate:CH_3OH:NH_3) UV and phosphomolybdic acid; υ_{max} (Film, cm^{-1}) 2360 (w), 1696 (m), 1664 (s), 1604 (w), 1550 (w), 1449 (m), 1415 (m), 1336 (m), 1273 (w), 1204 (s), 1176 (s), 1136 (s), 1082 (m), 1050 (m); δ_H (300 MHz, $CDCl_3$) 8.90 (2H, d, *J*=7.4 Hz, $\mathbf{H}^{4/5}$), 7.70 (2H, d, *J*=7.4 Hz, $\mathbf{H}^{1/8}$), 7.45 (2H, t, *J*=7.4 Hz, $\mathbf{H}^{3/6}$), 7.35 (2H, t, *J*=7.4 Hz, $\mathbf{H}^{2/7}$), 4.45 (2H, d, *J*=6.6 Hz, $C\mathbf{H}_2CH^9$), 4.35 (1H, t, *J*=6.6 Hz, $C\mathbf{H}^9$), 3.35 (2H, br s, $NH_2CH_2C\mathbf{H}_2NH$), 2.30 (2H, br s, $NHCH_2C\mathbf{H}_2NH_2$); δ_C (75 MHz, $CDCl_3$) 156.44 (0), 143.91 (0), 140.85 (0), 127.74 (1), 127.16 (1), 125.18 (1), 120.26 (1), 65.59 (2), 46.73 (1), 38.73 (2), 38.12 (2); LRMS (ES+) *m/z* 283.3 $(M+H)^+$ (100 %); HRMS (ES+) *m/z* 283.1438 $(M+H)^+$; $C_{17}H_{19}N_2O_2$ requires 283.1441.

tert-Butyl 2-[(2-[(9*H*-9-fluorenylmethoxy)carbonyl]aminoethyl)amino]acetate (22) and _tert_-butyl 2-[[2-(_tert_-butoxy)-2-oxoethyl](2-[(9*H*-9-fluorenylmethoxy)carbonyl]aminoethyl)amino]acetate (23)

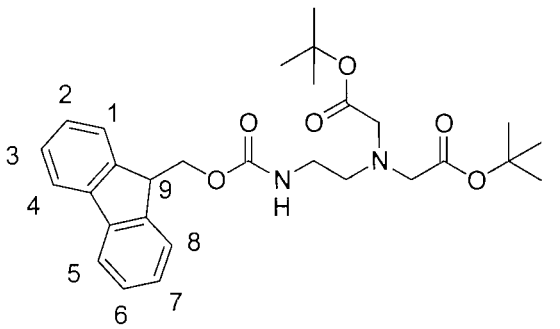
tert-Butyl bromoacetate (0.26 mL, 1.77 mmol) in DMF (20 mL) was added over 4 hours to a rapidly stirred solution of amine **(20)** (0.50 g, 1.77 mmol) and DIPEA (0.30 mL, 1.77 mmol) in DMF (5 mL) at 0 °C. The mixture was allowed to warm to room temperature and stirred for 1 hour. The mixture was concentrated *in vacuo* and the residue dissolved in ethyl acetate (25 mL) and washed with water (3 x 25 mL). The organic layer was dried over Na_2SO_4, filtered and concentrated *in vacuo*. The residue was purified by column chromatography (ethyl acetate to 95:5 ethyl acetate:CH_3OH). Fractions containing the products were concentrated *in vacuo* to yield compounds **(22)** (0.20 g, 28 %) and **(23)** (0.31 g, 35 %) as colourless solids.

Data for *tert*-butyl 2-[(2-[(9*H*-9-fluorenylmethoxy)carbonyl]aminoethyl)amino]acetate hydrochloride (22)



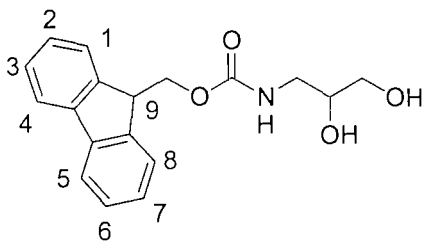
Data: R_f = 0.55 (95:5 ethyl acetate:CH_3OH) UV and phosphomolybdic acid; υ_{max} (Film, cm^{-1}) 2978 (w), 2934 (w), 2360 (w), 2338 (w), 1719 (s), 1691 (s), 1532 (w), 1477 (w), 1450 (m), 1395 (w), 1369 (m), 1249 (m), 1205 (s), 1154 (s), 1032 (w); δ_H (300 MHz, CD_3OD) 7.05 (2H, d, *J*=7.5 Hz, $\mathbf{H}^{4/5}$), 6.95 (2H, d, *J*=7.5 Hz, $\mathbf{H}^{1/8}$), 6.65 (2H, t, *J*=7.5 Hz, $\mathbf{H}^{3/6}$), 6.60 (2H, t, *J*=7.5 Hz, $\mathbf{H}^{2/7}$), 3.65 (2H, d, *J*=6.5 Hz, $C\mathbf{H}_2CH^9$), 3.50 (1H, t, *J*=6.5 Hz, $C\mathbf{H}^9$), 2.65 (2H, s, $NHC\mathbf{H}_2CO_2C(CH_3)_3$), 2.55 (2H, t, *J*=5.5 Hz, $NHC\mathbf{H}_2CH_2NH$), 2.05 (2H, t, *J*=5.5 Hz, $NHCH_2C\mathbf{H}_2NH$), 0.75 (9H, s, $C(C\mathbf{H}_3)_3$); δ_C (75 MHz, CD_3OD) 171.55 (0), 158.98 (0), 145.22 (0), 142.52 (0), 128.70 (1), 128.06 (1), 126.09 (1), 120.85 (1), 82.84 (0), 67.69 (2), 51.14 (2), 49.40 (2), 48.36 (1), 40.68 (2), 28.22 (3); LRMS (ES+) *m/z* 419.4 $(M+Na)^+$ (10 %), 397.3 $(M+H)^+$ (100 %); HRMS (ES+) *m/z* 397.2120 $(M+H)^+$; $C_{23}H_{29}N_2O_4$ requires 397.2122.

Data for *tert*-butyl 2-[[2-(*tert*-butoxy)-2-oxoethyl](2-[(9*H*-9-fluorenylmethoxy)carbonyl]aminoethyl)amino]acetate (23)



Data: R_f = 0.94 (95:5 ethyl acetate:CH_3OH) UV and phosphomolybdic acid; υ_{max} (Film, cm^{-1}) 3338 (w), 2974 (w), 2361 1718 (s), 1509 (m), 1477 (w), 1450 (m), 1449 (m), 1392 (w), 1367 (m), 1258 (m), 1220 (m), 1142 (m) 1078 (m), 1052 (m), 1022 (m); δ_H (300 MHz, d_6-DMSO) 8.00 (2H, d, *J*=7.4 Hz, $\mathbf{H}^{4/5}$), 7.75 (2H, d, *J*=7.4 Hz, $\mathbf{H}^{1/8}$), 7.50 (2H, t, *J*=7.4 Hz, $\mathbf{H}^{3/6}$), 7.40 (2H, t, *J*=7.4 Hz, $\mathbf{H}^{2/7}$), 4.45 (2H, d, *J*=6.5 Hz, $C\mathbf{H}_2CH^9$), 4.35 (1H, t, *J*=6.5 Hz, $C\mathbf{H}^9$), 3.45 (4H, s, $NC\mathbf{H}_2CO_2C(CH_3)_3$), 3.30 (2H, t, *J*=5.7 Hz, $NHC\mathbf{H}_2CH_2N$), 2.75 (2H, t, *J*=5.7 Hz, $NHCH_2C\mathbf{H}_2N$), 1.50 (18H, s, $C(C\mathbf{H}_3)_3$); δ_C (75 MHz, d_6-DMSO) 170.16 (0), 155.81 (0), 143.74 (0), 140.57 (0), 127.43 (1), 127.12 (1), 124.96 (1), 119.94 (1), 80.10 (0), 65.16 (2), 55.44 (2), 52.85 (2), 46.57 (1), 39.15 (2), 27.61 (3); LRMS (ES+) *m/z* 511.4 $(M+H)^+$ (100 %), 533.3 $(M+Na)^+$ (65 %); HRMS (ES+) *m/z* 511.2813 $(M+H)^+$; $C_{29}H_{39}N_2O_6$ requires 511.2802.

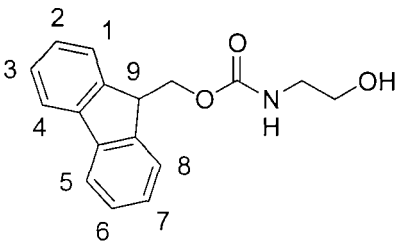
9*H*-9-Fluorenylmethyl *N*-(2,3-dihydroxypropyl)carbamate (25)



Fmoc-Su (1.11 g, 3.29 mmol) in CH_2Cl_2 (5 mL) was added to a solution of 3-amino-1,2-propanediol **(24)** (0.26 mL, 3.29 mmol) and DIPEA (0.62 mL, 3.29 mmol) in CH_2Cl_2 (2 mL). The reaction mixture was stirred overnight under an atmosphere of argon. The reaction mixture was washed with 10 % citric acid (3 x 10 mL) and saturated KCl (2 x 10 mL). The organic layer was dried over Na_2SO_4, filtered and concentrated *in vacuo.* The resulting residue was triturated with Et_2O yielding diol **(25)** as a colourless solid, which was isolated by filtration and dried in air (0.63 g, 61 %).

Data: R_f = 0.25 (ethyl acetate) UV and ninhydrin; υ_{max} (Solid, cm^{-1}) 3320 (m), 2949 (w), 2357 (w), 1685 (s), 1551 (s), 1438 (m), 1260 (s), 1158 (m), 1099 (m), 1012 (w); δ_H (400 MHz, d_6-DMSO) (1:1 rotamers) 7.95 (4H, d, *J*=7.4 Hz, $\mathbf{H}^{4/5}$), 6.81 (4H, d, *J*=7.4 Hz, $\mathbf{H}^{1/8}$), 7.53 (4H, t, *J*=7.4 Hz, $\mathbf{H}^{3/6}$), 7.44 (4H, t, *J*=7.4 Hz, $\mathbf{H}^{2/7}$), 7.28 (2H, br s, N**H**), 4.77 (2H, m, C$\mathbf{H}_2$OH), 4.59 (2H, m, C$\mathbf{H}_2$OH), 4.38 (4H, d, *J*=6.5 Hz, C$\mathbf{H}_2$CH^9), 4.32 (2H, t, *J*=6.5 Hz, C**H**^{9}), 3.60 (1H, m, NHCH$_2$C**H**OH), 3.42 (1H, m, NHCH$_2$C**H**OH), 3.22 (2H, m, NHC$\mathbf{H}_2$), 3.04 (2H, m, NHC$\mathbf{H}_2$); δ_C (75 MHz, d_6-DMSO) 159.10 (0), 145.22 (0), 142.51 (0), 128.68 (1), 128.05 (1), 126.07 (1), 120.83 (1), 72.16 (2), 67.68 (1), 64.90 (2), 48.35 (1), 44.52 (2); LRMS (ES+) *m/z* 314.2 $(M+H)^+$ (20%), 336.2 $(M+Na)^+$ (100%); HRMS (ES+) *m/z* 336.1201 $(M+Na)^+$; $C_{18}H_{19}NO_4Na$ requires 336.1206.

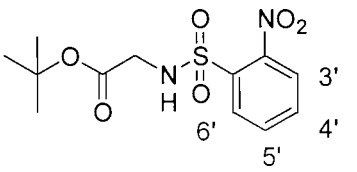
9*H*-9-Fluorenylmethyl *N*-(2-hydroxyethyl)carbamate (28)



Following the procedure described by Rodriguez *et al.*[173] *iso*-Butyl chloroformate (0.86 mL, 5.10 mmol) was added to a solution of FmocGlyOH **(27)** (1.50 g, 5.10 mmol) and NMM (0.73 mL, 5.10 mmol) in 1,2-dimethoxyethane (DME) (10 mL) at -15 °C. The reaction mixture was stirred for 5 minutes and filtered. $NaBH_4$ (0.35 g, 7.57 mmol) in water (5 mL) was added to the filtrate at -15 °C. The reaction mixture was stirred for a further 10 minutes and water (200 mL) was added. The resulting solid was isolated by filtration and recrystallised in CH_2Cl_2/hexane to give alcohol **(28)** as a colourless solid (1.21 g, 85 %).

Data: R_f = 0.39 (95:5 in CH_2Cl_2:CH_3OH) UV and ninhydrin; υ_{max} (Solid, cm^{-1}) 3465 (m), 3336 (s), 3072 (w), 2943 (m), 2884 (w), 1685 (s), 1540 (s), 1470 (w), 1448 (m), 1411 (w), 1357 (w), 1271 (s), 1153 (m), 1099 (m), 1072 (m), 1018 (m); δ_H (300 MHz, $CDCl_3$) 7.80 (2H, d, *J*=7.5 Hz, $\mathbf{H}^{4/5}$), 7.62 (2H, d, *J*=7.5 Hz, $\mathbf{H}^{1/8}$), 7.42 (2H, t, *J*=7.5 Hz, $\mathbf{H}^{3/6}$), 7.31 (2H, t, *J*=7.5 Hz, $\mathbf{H}^{2/7}$), 5.17 (1H, br s, N**H**), 4.44 (2H, d, *J*=6.4 Hz, $C\mathbf{H}_2CH^9$), 4.23 (1H, t, *J*=6.4 Hz, $C\mathbf{H}^9$), 3.74 (2H, t, *J*=5.8 Hz, $NHCH_2C\mathbf{H}_2OH$), 3.38 (2H, t, *J*=5.8 Hz, $NHC\mathbf{H}_2CH_2OH$); δ_C (75 MHz, $CDCl_3$) 164.03 (0), 144.01 (0), 141.48 (0), 127.87 (1), 127.21 (1), 125.16 (1), 120.15 (1), 66.91 (2), 62.50 (2), 47.38 (1), 43.60 (2); LRMS (ES+) *m/z* 306.2 $(M+Na)^+$ (100%); HRMS (ES+) *m/z* 306.1095 $(M+Na)^+$; $C_{17}H_{17}NO_3Na$ requires 306.1101.

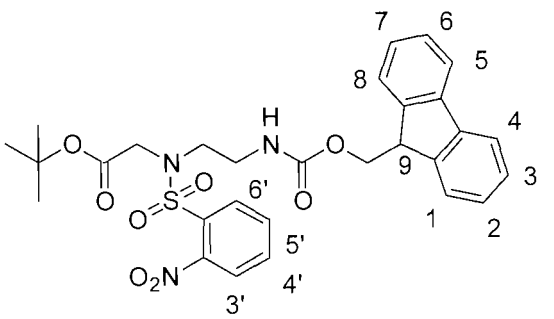
tert-Butyl 2-[(2'-nitrophenyl)sulfonyl]aminoacetate (30)

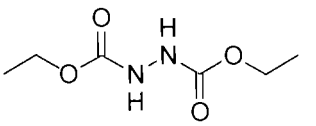


Following a method adapted from the procedure described by Fukuyama *et al.*[167] 2-Nitrosulfonyl chloride (0.68 g, 3.05 mmol) in CH_2Cl_2 (5 mL) was added drop wise to *tert*-butyl 2-aminoacetate•HCl **(29)** (0.40 g, 3.05 mmol) in CH_2Cl_2 (15 mL) adjusted to pH 8 with Et_3N. The reaction mixture was stirred for 1 hour at room temperature. The reaction mixture was diluted with CH_2Cl_2 (20 mL) and washed with 10 % citric acid (50 mL) and saturated KCl (50 mL). The organic layer was dried over Na_2SO_4, filtered and concentrated *in vacuo.* The resulting yellow residue was recrystallised in ethyl acetate/hexane to yield sulfonamide **(30)** as a colourless solid (0.82 g, 85 %).

Data: R_f = 0.45 (CH_2Cl_2) UV and ninhydrin; υ_{max} (Film, cm^{-1}) 3346 (br w), 3095 (w), 2979 (w), 1734 (s), 1593 (w), 1536 (s), 1441 (m), 1415 (m), 1354 (s), 1303 (m), 1244 (m), 1175 (s), 1122 (s), 1058 (w), 1001 (w); δ_H (400 MHz, $CDCl_3$) 8.13 (1H, m, $\mathbf{H}^{3'}$), 7.97 (1H, m, $\mathbf{H}^{6'}$), 7.77 (2H, m, $\mathbf{H}^{5'/4'}$), 6.07 (1H, br t, *J*=5.7 Hz, N**H**), 3.93 (2H, d, *J*=5.7 Hz, NHC$\mathbf{H}_2$), 1.38 (9H, s, C(C$\mathbf{H}_3$)$_3$); δ_C (100 MHz, $CDCl_3$) 167.82 (0), 148.31 (0), 134.37 (0), 134.06 (1), 133.25 (1), 30.99 (1), 126.06 (1), 83.38 (0), 46.02 (2), 38.21 (3); LRMS (ES+) *m/z* 339.1 $(M+Na)^+$ (100%); (ES-) 315.2 $(M-H)^-$ (100 %); HRMS (ES+) *m/z* 655.1353 $(2M+Na)^+$; $C_{24}H_{32}N_4O_{12}S_2Na$ requires 655.1350.

tert-Butyl 2-[(2-[(9*H*-9-fluorenylmethoxy)carbonyl]aminoethyl)-2'-nitroanilino]acetate (31) and diethyl 1,2-hydrazinedicarboxylate (65)

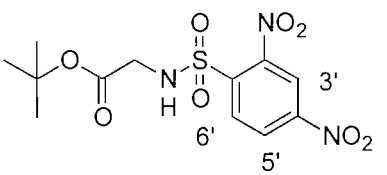




Following a method adapted from the procedure described by Fukuyama *et al.*[167] Diethyl azodicarboxylate (0.21 mL, 1.32 mmol) in anhydrous CH_2Cl_2 (5 mL) was added to a solution of sulfonamine **(30)** (0.28 g, 1.32 mmol), triphenyl phosphine (0.35 g, 1.88 mmol) and alcohol **(28)** (0.25 g, 0.88 mmol) in anhydrous CH_2Cl_2 (10 mL). The reaction mixture was stirred for 30 minutes at room temperature under an atmosphere of argon. The reaction mixture was concentrated *in vacuo* and the resulting residue was dissolved in Et_2O. On cooling a colourless solid formed which was isolated by filtration. The filtrate was concentrated *in vacuo* and purified by silica gel column chromatography (2:1 hexane:ethyl acetate). Fractions containing the products were concentrated *in vacuo* to yield a 1:1 mixture of diethyl 1,2-hydrazinedicarboxylate **(65)** and sulfonamide **(31)** (0.73g, 71 %).

Data: R_f = 0.28 (1:1 ethyl acetate:hexane) UV, δ_H (400 MHz, $CDCl_3$) 7.99 (1H, d, *J*=5.0 Hz, **H**$^{3'}$), 7.70 (2H, d, *J*=7.5 Hz, **H**$^{4/5}$), 7.53-7.60 (2H, m, **H**$^{4'/5'}$), 7.53 (2H, d, *J*=7.5 Hz, **H**$^{1/8}$), 7.48 (1H, d, *J*=5.0 Hz, **H**$^{6'}$), 7.33 (2H, t, *J*=7.5 Hz, **H**$^{3/6}$), 7.22 (2H, d, *J*=7.5 Hz, **H**$^{2/7}$), 6.38 (2H, br s, N**H**N**H**), 5.42-5.47 (1H, m, N**H**), 4.25 (2H, d, *J*=7.0 Hz, C**H**$_2$CH^9), 4.15 (4H, q, *J*=7.0 Hz, C**H**$_2$CH$_3$), 4.13 (1H, t, *J*=7.0 Hz, C**H**^{9}), 4.00 (2H, s, NC**H**$_2$CO$_2$C(CH$_3$)$_3$), 3.45 (2H, t, *J*=5.5 Hz, NC**H**$_2$CH$_2$NH), 3.34 (2H, t, *J*=5.5 Hz, NCH$_2$C**H**$_2$NH), 1.34 (9H, s, (C**H**$_3$)$_3$), 1.20 (6H, t, *J*=7.0 Hz, C**H**$_3$); δ_C (100 MHz, $CDCl_3$) 168.61 (0), 156.96 (0), 149.32 (0), 144.36 (0), 141.68 (0), 134.06 (0), 133.42 (0), 132.19 (1), 131.60 (1), 129.30 (1), 128.10 (1), 127.50 (1), 125.66 (1), 124.58 (1), 120.35 (1), 83.22 (0), 67.38 (2), 62.71 (2), 50.00 (2), 49.08 (2), 47.54 (1), 39.24 (2), 28.32 (3), 14.83 (3); LRMS (ES+) *m/z* 199.2 $(M+Na)^+$ (75 %), 604.2 $(M+Na)^+$ (100%); HRMS (ES+) *m/z* 604.1728 $(M+Na)^+$; $C_{29}H_{31}N_3O_8SNa$ requires 604.1724.

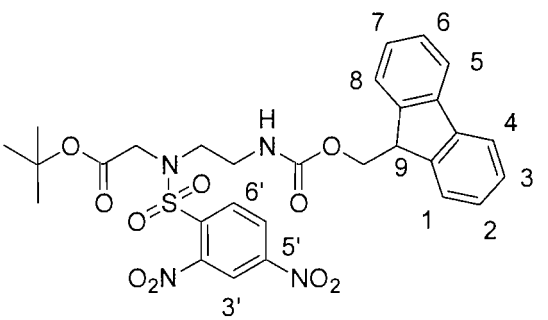
tert-Butyl 2-[(2',4'-dinitrophenyl)sulfonyl]aminoacetate (32)



Following the procedure described by Fukuyama *et al.*[168] 2,4-Dinitrosulfonyl chloride (5.54 g, 20.79 mmol) in CH_2Cl_2 (30 mL) was added to a solution of *tert*-butyl 2-aminoacetate•HCl **(29)** (3.00 g, 20.79 mmol) in CH_2Cl_2 (40 mL) at -15 °C (adjusted to pH 8 with lutidine). The reaction mixture was allowed to warm to room temperature and stirred for 1 hour. The reaction mixture was diluted with CH_2Cl_2 (30 mL) and washed with 10 % citric acid (100 mL) then saturated KCl (100 mL). The organic layer was dried over Na_2SO_4, filtered and concentrated *in vacuo*. The resulting yellow residue was recrystallised from ethyl acetate/hexane to yield sulfonamide **(32)** as an orange crystalline solid (6.83 g, 91 %).

Data: R_f = 0.28 (CH_2Cl_2) UV; MP = 123-125 °C; υ_{max} (film, cm^{-1}) 3350 (w), 3112 (w), 2982 (w), 1739 (m), 1605 (w), 1540 (s), 1413 (w), 1350 (s), 1251 (m), 1156 (s), 1116 (w); (δ_H 400 MHz, $CDCl_3$) 8.68 (1H, d, *J*=2.0 Hz, $\mathbf{H}^{3'}$), 8.45 (1H, dd, *J*=2.0, 8.5 Hz, $\mathbf{H}^{5'}$), 8.23 (1H, d, *J*=8.5 Hz, $\mathbf{H}^{6'}$), 3.89 (2H, s, $C\mathbf{H}_2NH$), 1.28 (9H, s, $C(C\mathbf{H}_3)_3$); δ_C (100 MHz, $CDCl_3$) 167.95 (0), 150.17 (0), 148.40 (0), 140.24 (0), 132.50 (1), 127.40 (1), 121.42 (1), 83.81 (0), 46.11 (2), 28.23 (3); LRMS (ES+) *m/z* 384.0 $(M+Na)^+$ (100%); (ES-) *m/z* 360.0 $(M-H)^-$ (100%); HRMS (ES+) *m/z* 745.1052 $(2M+Na)^+$; $C_{24}H_{30}N_6O_{16}S_2Na$ requires 745.1052.

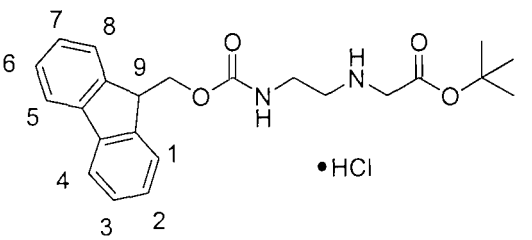
tert-Butyl 2-[[(2',4'-dinitrophenyl)sulfonyl](2-[(9*H*-9-fluorenylmethoxy)carbonyl]aminoethyl)amino]acetate (33)



Following the procedure described by Fukuyama *et al.*[168] Diisopropylazodicarboxylate (1.68 mL, 8.30 mmol) in anhydrous CH_2Cl_2 (5 mL) was added to a solution of sulfonamide **(32)** (2.00 g, 5.54 mmol), triphenyl phosphine (2.18 g, 8.30 mmol) and alcohol **(28)** (1.57 g, 5.54 mmol) in anhydrous CH_2Cl_2 (10 mL). The reaction mixture was stirred overnight at room temperature under an atmosphere of argon. The reaction mixture was pre-absorbed directly on to silica gel and purified by silica gel column chromatography (90:10 hexane:ethylacetate) to yield sulfonamide **(33)** as a yellow solid (2.86 g, 82 %).

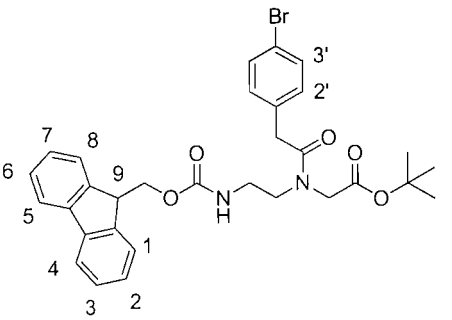
Data: R_f = 0.34 (1:1 ethyl acetate:hexane) UV and phophomolybdic acid; υ_{max} (film, cm^{-1}) 1719 (m), 1538 (s), 1453 (w), 1369 (s), 1235 (s), 1685 (s), 1103 (w), 1017 (w); δ_H (400 MHz, $CDCl_3$), 8.36 (1H, dd, *J*=2.0, 8.5 Hz, $\mathbf{H}^{5'}$), 8.31 (1H, d, *J*=2.0 Hz, $\mathbf{H}^{3'}$), 8.18 (1H, d, *J*=8.5 Hz, $\mathbf{H}^{6'}$), 7.69 (2H, d, *J*=7.5 Hz, $\mathbf{H}^{4/5}$), 7.51 (2H, d, *J*=7.5 Hz, $\mathbf{H}^{1/8}$), 7.32 (2H, t, *J*=7.5 Hz, $\mathbf{H}^{3/6}$), 7.24 (2H, t, *J*=7.5 Hz, $\mathbf{H}^{2/7}$), 5.27 (1H, br s, N**H**), 4.28 (2H, d, *J*=7.5 Hz, $C\mathbf{H}_2CH^9$), 4.13 (1H, t, *J*=7.5 Hz, $C\mathbf{H}^9$), 3.89 (2H, s, $COC\mathbf{H}_2N$), 3.46 (2H, t, *J*=6.0 Hz, $NC\mathbf{H}_2CH_2NH$), 3.35 (2H, t, *J*=6.0 Hz, $NCH_2C\mathbf{H}_2NH$), 1.33 (9H, s, $C(C\mathbf{H}_3)_3$); δ_C (100 MHz, $CDCl_3$) 169.48 (0), 158.23 (0), 151.36 (0), 149.72 (0), 145.50 (0), 142.99 (0), 140.09 (0), 134.39 (1), 129.43 (1), 128.77 (1), 127.76 (1), 126.80 (1), 121.70 (1), 121.23 (1), 84.98 (0), 68.74 (2), 51.19 (2), 50.51 (2), 48.84 (1), 40.45 (2), 29.60 (3); LRMS (ES+) *m/z* 649.3 $(M+Na)^+$ (100%); HRMS (ES+) *m/z* 649.1552 $(M+Na)^+$; $C_{29}H_{30}N_4O_{10}SNa$ requires 649.1578.

***tert*-Butyl 2-[(2-[(9*H*-9-fluorenylmethoxy)carbonyl]aminoethyl)amino]acetate hydrochloride (22)**



Following the procedure described by Fukuyama *et al.*[168] Mercaptoacetic acid (0.95 g, 10.31 mmol) and DIPEA (2.6 mL, 15.46 mmol) in CH_2Cl_2 (5 mL) was added to a solution of sulfonamide **(33)** (3.23 g, 5.15 mmol) in CH_2Cl_2 (20 mL). The reaction mixture was stirred for 20 minutes, diluted with CH_2Cl_2 (25 mL) and extracted with saturated $NaHCO_3$ (50 mL) followed by saturated KCl (50 mL) until the aqueous layer was clear. The organic layer was extracted with 1 M HCl (50 mL), dried over Na_2SO_4, filtered and partially concentrated *in vacuo* and stored for 2 days at – 15 °C yielding PNA backbone **(22)** as the hydrochloride salt (1.89 g, 92 %).

tert-Butyl 2-[[2-(4'-bromophenyl)acetyl](2-[(9*H*-9-fluorenylmethoxy)carbonyl]aminoethyl)amino]acetate (34)

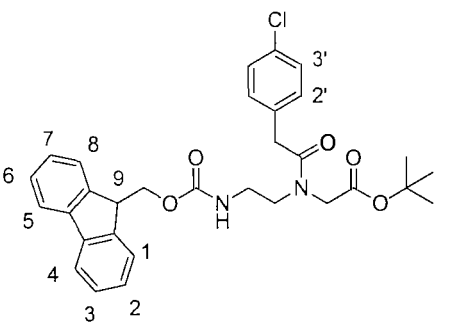


4-Bromophenylacetic acid (0.57 g, 2.65 mmol) was added to a solution of PNA backbone **(22)** (1.00 g, 2.52 mmol) and EDC (0.73 g, 3.78 mmol) (adjusted to pH 8 with DIPEA) in anhydrous DMF (30 mL). The reaction mixture was stirred for 2 hours at room temperature under an atmosphere of argon. The reaction mixture was then diluted with CH_2Cl_2 (150 ml) and washed with saturated KCl (3 x 100 mL). The organic phase was dried over Na_2SO_4, filtered, pre-absorbed directly on to silica gel and purified by silica gel column chromatography (1:1 ethyl acetate:hexane). Fractions containing the product were concentrated *in vacuo* to yield monomer **(34)** as a colourless foam (0.81 g, 54 %).

Data: R_f = 0.62 (99:1 ethyl acetate:NH_3) UV and phosphomolybdic acid; υ_{max} (film, cm^{-1}) 3319 (br w), 3065 (w), 2979 (w), 2938 (w), 1718 (s), 1645 (s), 1518 (m), 1488 (m), 1449 (s), 1415 (m), 1405 (w), 1367 (m), 1318 (w), 1243 (s), 1180 (s), 1104 (w), 1071 (w), 1012 (m); δ_H (400 MHz, $CDCl_3$) (2:1 rotamers) 7.69 (6H, d, *J*=7.0 Hz, $\mathbf{H}^{4/5}$), 7.53-7.50 (6H, m, $\mathbf{H}^{1/8}$), 7.30-7.36 (12H, m, $\mathbf{H}^{3'}$, $\mathbf{H}^{3/6}$), 7.21-7.25 (6H, m, $\mathbf{H}^{2/7}$), 7.00-7.04 (6H, m, $\mathbf{H}^{2'}$), 5.76 (2H, br s, N**H**), 5.39 (1H, br s, N**H**), 4.32-4.36 (4H, m, $C\mathbf{H}_2CH^9$), 4.25-4.28 (2H, m, $C\mathbf{H}_2CH^9$), 4.10-4.16 (3H, m, $C\mathbf{H}^9$), 3.88 (4H, s, $NC\mathbf{H}_2CO_2C(CH_3)_3$), 3.82 (2H, s, $NC\mathbf{H}_2CO_2C(CH_3)_3$), 3.56 (4H, s, $COC\mathbf{H}_2Ar$), 3.47-3.52 (6H, m, $NHCH_2C\mathbf{H}_2N$, $COC\mathbf{H}_2Ar$), 3.37-3.43 (2H, m, $NHCH_2C\mathbf{H}_2N$), 3.26-3.35 (6H, m, $NHC\mathbf{H}_2CH_2N$, $NHC\mathbf{H}_2CH_2N$), 1.46 (18H, s, $C(C\mathbf{H}_3)_3$), 1.41 (9H, s, $C(C\mathbf{H}_3)_3$);

δ_C (100 MHz, $CDCl_3$) 172.70 (0), 172.06 (0), 170.258 (0), 169.33 (0), 157.44 (0), 144.79 (0), 144.63 (0), 142.16 (0), 134.45 (0), 134.13 (0), 132.62 (0), 132.54 (0), 131.60 (1), 131.56 (1), 128.60 (1), 128.51 (1), 127.94 (1), 127.89 (1), 125.96 (1), 125.86 (1), 121.90 (1), 121.71 (1), 120.86 (1), 120.80 (1), 84.00 (0), 83.20 (0), 67.80 (2), 67.62 (2), 61.21 (2), 52.46 (2), 50.75 (2), 50.45 (2), 49.04 (2), 48.08 (1), 48.03 (1), 40.79 (2), 40.25 (2), 28.92 (3), 28.86 (3); LRMS (ES+) *m/z* 615.5, 617.2 $(M+Na)^+$ (Br isotope pattern 100 %); HRMS (ES+) *m/z* 615.1454 $(M+Na)^+$; $C_{31}H_{33}N_2O_5Br^{79}Na$ requires 615.1454.

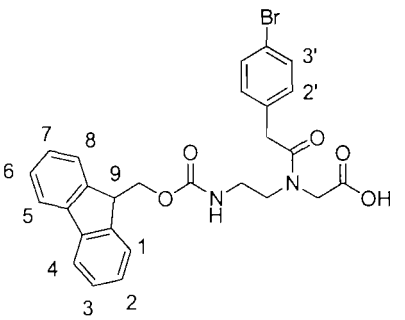
tert-Butyl 2-[[2-(4'-chlorophenyl)acetyl](2-[(9*H*-9-fluorenylmethoxy)carbonyl]aminoethyl)amino]acetate (35)



4-Chlorophenylacetic acid (0.79 g, 4.63 mmol) was added to a solution of PNA backbone **(22)** (1.67 g, 4.21 mmol) and EDC (1.21 g, 6.32 mmol) (adjusted to pH 8 with DIPEA) in anhydrous DMF (30 mL). The reaction mixture was stirred for 60 minutes at room temperature under an atmosphere of argon. The reaction mixture was then diluted with CH_2Cl_2 (150 mL) and extracted with saturated KCl (3 x 100 mL). The organic phase was dried over Na_2SO_4, filtered, pre-absorbed on to silica gel and purified by silica gel column chromatography (1:1 ethyl acetate:hexane). Fractions containing the product were concentrated *in vacuo* to yield monomer **(35)** as a colourless foam (1.56 g, 67 %).

Data: R_f = 0.64 (95:5 ethyl acetate:CH_3OH) UV and phosphomolybdic acid; υ_{max} (film, cm^{-1}) 3311 (br w), 2977 (w), 2941 (w), 1720 (s), 1646 (s), 1515 (m), 1492 (m), 1449 (s), 1407 (w), 1367 (m), 1318 (w), 1231 (s), 1152 (s), 1090 (w), 1015 (w); δ_H (400 MHz, $CDCl_3$) (2:1 rotamers) 7.67-7.71 (6H, m, **H**$^{4/5}$), 7.48-7.53 (6H, m, **H**$^{1/8}$), 7.29-7.35 (6H, m, **H**$^{3/6}$), 7.19-7.24 (6H, m, **H**$^{2/7}$), 7.17 (6H, dd, *J*=2.0 ,9.1, **H**$^{3'}$), 7.08 (6H, dd, *J*=2.0 ,9.1, **H**$^{2'}$), 5.77 (2H, br t, N**H**), 5.39 (1H, br t, N**H**), 4.30-4.36 (4H, m, C**H**$_2$CH^9), 4.25-4.28 (2H, m, C**H**$_2$CH^9), 4.12-4.18 (3H, m, C**H**^{9}), 3.88 (4H, s, NC**H**$_2$CO$_2$C(CH$_3$)$_3$), 3.83 (2H, s, NC**H**$_2$CO$_2$C(CH$_3$)$_3$), 3.58 (4H, s, COC**H**$_2$Ar), 3.48-3.53 (6H, m, NHCH$_2$C**H**$_2$N, COC**H**$_2$Ar), 3.38-3.45 (2H, m, NHCH$_2$C**H**$_2$N), 3.27-3.35 (4H, m, NHC**H**$_2$CH$_2$N), 3.23-3.27 (2H, m, NHC**H**$_2$CH$_2$N), 1.42 (18H, s, C(C**H**$_3$)$_3$), 1.39 (9H, s, C(C**H**$_3$)$_3$); δ_C (100 MHz, $CDCl_3$) 170.45 (0), 169.80 (0), 167.91 (0), 166.99 (0), 155.10 (0), 142.44 (0), 142.29 (0), 139.81 (0), 131.58 (0), 131.48 (0), 131.28 (0), 128.88 (1), 128.86 (1), 127.31 (1), 127.24 (1), 126.25 (1), 126.17 (1), 125.59 (1), 125.54 (1), 123.62 (1), 123.52 (1), 118.51 (1), 118.44 (1), 81.64 (0), 80.85 (0), 65.45 (2), 65.28 (2), 58.86 (1), 50.11 (2), 48.41 (2), 48.11 (2), 46.70 (2), 45.74 (2), 45.69 (2), 38.37 (2), 37.99 (2), 26.52 (3), 26.48 (3); LRMS (ES+) *m/z* 571.3, 573.3 $(M+Na)^+$ (Cl isotope pattern 100 %); HRMS (ES+) *m/z* 571.1970 $(M+Na)^+$; $C_{31}H_{33}N_2O_5Cl^{35}Na$ requires 571.1970.

2-[[2-(4'-Bromophenyl)acetyl](2-[(9*H*-9-fluorenylmethoxy)carbonyl]aminoethyl)amino]acetic acid (36)

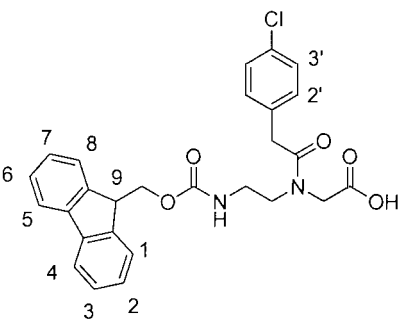


TFA (10 mL) was added to a solution of monomer **(34)** (1.00 g, 2.52 mmol) in CH_2Cl_2 (10 mL). The reaction mixture was stirred for 30 minutes at room temperature. The reaction mixture was concentrated *in vacuo* and Et_2O (10 mL) was added. The precipitate formed on cooling was washed with Et_2O (10 mL), dried *in vacuo* to yield monomer **(36)** as a colourless solid (0.81 g, 54 %).

Data: R_f = 0.42 (95:5 CH_2Cl_2:CH_3OH) UV and phosphomolybdic acid; υ_{max} (solid, cm^{-1}) 3286 (m), 3795 (w), 2891 (w), 17213 (s), 1672 (m), 1632 (s), 1592 (w), 1472 (m), 1445 (s), 1403 (m), 1360 (w), 1328 (m), 1275 (w), 1223 (m), 1191 (m), 1134 (m), 1095 (m), 1064 (m), 1009 (w); MP = 168-170 °C; δ_H (400 MHz, d_6-DMSO) (2:1 rotamers) 12.85 (3H, s, $CO_2\mathbf{H}$), 7.98-8.03 (6H, m, $\mathbf{H}^{4/5}$), 7.75-7.79 (6H, m, $\mathbf{H}^{1/8}$), 7.48-7.60 (14H, m, $\mathbf{H}^{3'}$, $\mathbf{H}^{3/6}$, N**H**), 7.40-7.46 (6H, m, $\mathbf{H}^{2/7}$), 7.28 (1H, br t, N**H**), 7.24-7.29 (6H, m, $\mathbf{H}^{2'}$), 4.42-4.46 (4H, m, $C\mathbf{H}_2CH^9$), 4.37-4.40 (2H, m, $C\mathbf{H}_2CH^9$), 4.30-4.36 (3H, m, $C\mathbf{H}^9$), 4.29 (4H, s, $NC\mathbf{H}_2COOH$), 4.06 (2H, s, $NC\mathbf{H}_2CO_2H$), 3.80 (4H, s, $COC\mathbf{H}_2Ar$), 3.68 (2H, s, $COC\mathbf{H}_2Ar$), 3.35-3.56 (6H, m, $NHCH_2C\mathbf{H}_2N$, $NHCH_2C\mathbf{H}_2N$), 3.24-3.34 (4H, m, $NHC\mathbf{H}_2CH_2N$), 3.20-3.23 (2H, m, $NHC\mathbf{H}_2CH_2N$); δ_C (100 MHz, d_6-DMSO) 171.60 (0), 171.25 (0), 171.17 (0), 170.80 (0), 156.80 (0), 156.58 (0), 144.30 (0), 141.22 (0), 135.62 (0), 135.55 (0), 132.10 (1), 131.91 (1), 131.42 (1), 131.35 (1), 128.07 (1), 127.51 (1), 125.59 (1), 125.51 (1), 120.58 (1), 119.98 (1), 65.90 (2), 48.30 (2), 47.94 (2), 47.19 (1), 47.02 (2), 39.28 (2), 38.87 (2), 38.44 (2);

LRMS (ES+) *m/z* 559.1, 561.1 $(M+Na)^+$ (Br isotope pattern 100 %); (ES-) *m/z* 535.2, 537.2 $(M-H)^-$ (Br isotope pattern 100 %); HRMS (ES+) *m/z* 559.0839 $(M+Na)^+$; $C_{27}H_{25}N_2O_5Br^{79}Na$ requires 559.0837; CHN C 59.68, H 4.67, N 5.16; $C_{27}H_{25}N_2O_5Br$ requires C 60.34, H 4.69, N 5.21.

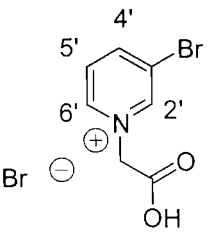
2-[[2-(4'-Chlorophenyl)acetyl](2-[(9*H*-9-fluorenylmethoxy)carbonyl]aminoethyl)amino]acetic acid (37)



TFA (10 mL) was added to a solution of monomer **(35)** (1.54 g, 2.80 mmol) in CH_2Cl_2 (10 mL). The reaction mixture was stirred for 60 minutes at room temperature. The reaction mixture was concentrated *in vacuo* and Et_2O (10 mL) was added. The precipitate formed on cooling was washed with Et_2O (10 mL) then dried *in vacuo* to yield monomer **(37)** as a colourless solid (1.32 g, 96 %).

Data: R_f = 0.44 (95:5 CH_2Cl_2:CH_3OH) UV and phosphomolybdic acid; υ_{max} (solid, cm^{-1}) 3282 (w), 2979 (w), 2949 (w), 2634 (w), 2565 (w), 1723 (s), 1673 (m), 1632 (s), 1596 (w), 1492 (w), 1471 (m), 1444 (s), 1404 (m), 1361 (w), 1329 (m), 1277 (w), 1224 (m), 1192 (m), 1134 (m), 1094 (m), 1064 (m), 1023 (w); δ_H (400 MHz, d_6-DMSO) (2:1 rotamers) 12.75 (3H, s, CO_2**H**), 7.98-8.01 (6H, m, **H**$^{4/5}$), 7.7-7.80 (6H, m, **H**$^{1/8}$), 7.50-7.54 (6H, m, **H**$^{3/6}$), 7.40-7.45 (14H, m, **H**$^{2/7}$, **H**$^{3'}$, N**H**), 7.28-7.34 (7H, m, **H**$^{2'}$, N**H**), 4.40-4.44 (4H, m, C**H**$_2$CH^9), 4.38-4.40 (2H, m, C**H**$_2$CH^9), 4.29-4.32 (3H, m, C**H**^{9}), 4.32 (4H, s, NC**H**$_2$$CO_2$H), 4.06 (2H, s, NC**H**$_2$$CO_2$H), 3.82 (4H, s, COC**H**$_2$Ar), 3.69 (2H, s, COC**H**$_2$Ar), 3.40-3.52 (6H, m, NHCH_2C**H**$_2$N, NHCH_2C**H**$_2$N), 3.26-3.31 (4H, m, NHC**H**$_2$$CH_2$N), 3.23-3.26 (2H, m, NHC**H**$_2$$CH_2$N); δ_C (100 MHz, d_6-DMSO) 170.72 (0), 170.45 (0), 170.30 (0), 169.99 (0), 155.92 (0), 155.70 (0), 143.42 (0), 140.33 (0), 134.31 (0), 134.23 (0), 130.82 (1), 130.62 (1), 127.62 (1), 127.54 (1), 127.19 (1), 126.63 (1), 124.70 (1), 124.63 (1), 119.70 (1), 65.01 (2), 49.66 (2), 47.42 (2), 47.06 (2), 46.31 (1), 46.15 (1), 39.72 (2), 39.52 (2), 39.31 (2), 37.92 (2), 39.72 (2), 37.50 (2); LRMS (ES+) *m/z* 515.1.1, 517.1 $(M+Na)^+$ (Cl isotope pattern 100 %); (ES-) 537.1.1, 539.1 $(M+2Na-H)^+$ (Cl isotope pattern 80 %), *m/z* 491.1, 493.1 $(M-H)^-$ (Cl isotope pattern 100 %); HRMS (ES+) *m/z* 515.1344 $(M+Na)^+$; $C_{27}H_{25}N_2O_5Cl^{35}Na$ requires 515.1351.

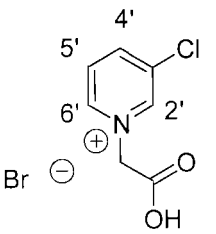
2-(3'-Bromo-1'-pyridiniumyl)acetic acid bromide (40)



Following the procedure described by Barry *et al.*[169] Bromoacetic acid (1.39 g, 10.00 mmol) was added to a solution of 3-bromopyridine **(38)** (1.580 g, 10.00 mmol) in anhydrous acetonitrile (30 mL). The reaction mixture was stirred for 2 hours under an atmosphere of argon. On cooling colourless crystals formed which were isolated by filtration and washed with ice-cold acetonitrile (2 x 10 mL) and Et_2O (2 x 10 mL). The crystals were dried *in vacuo* to yield acid **(40)** as a pale yellow solid (2.64 g, 90 %).

Data: R_f = 0.09 (8:2 CH_2Cl_2:CH_3OH) UV; υ_{max} (solid, cm^{-1}) 2840 (br w), 3683 (w), 2570 (w), 1731 (s), 1632 (m), 1625 (m), 1562 (w), 1487 (m), 1461 (w), 1418 (m), 1390 (m), 1330 (w), 1311 (w), 1182 (s), 1101 (m), 1031 (w); δ_H (400 MHz, D_2O) 9.14 (1H, s, $\mathbf{H}^{2'}$), 8.83 (2H, 2 x d, *J*=6.6, 8.0, Hz, $\mathbf{H}^{6'/4'}$), 8.03 (1H, dd, *J*=6.6, 8.0, Hz, $\mathbf{H}^{5'}$), 5.45 (2H, s, $C\mathbf{H}_2$); δ_C (100 MHz, D_2O) 169.02 (0), 149.68 (1), 147.32 (1), 144.95 (1), 128.95 (0), 123.00 (1), 61.93 (2); LRMS (ES+) *m/z* 216.0, 218.0 $(M)^+$ (Br isotope pattern) (100%), 238.0, 240.0 $(M+Na-H)^+$ (Br isotope pattern) (35 %); HRMS (ES+) *m/z* 215.9655 $(M)^+$; $C_7H_7NO_2Br^{79}$ requires 215.9655.

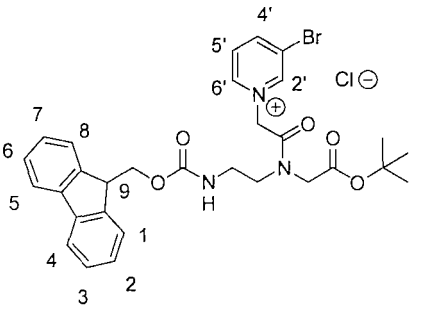
2-(3'-Chloro-1'-pyridiniumyl)acetic acid bromide (31)



Following a method adapted from the procedure described by Barry *et al.*[169] Bromoacetic acid (1.39 g, 10.00 mmol) was added to a solution of 3-chloropyridine **(39)** (1.13 g, 10.00 mmol) in anhydrous acetonitrile (20 mL). The reaction mixture was refluxed for 2 hours under an atmosphere of argon. On cooling yellow crystals formed which were isolated by filtration and washed with ice-cold acetonitrile (2 x 10 mL) and Et_2O (2 x 10 mL). The crystals were dried *in vacuo* to yield acid **(41)** as a yellow solid (2.13 g, 86 %).

Data: R_f = 0.10 (8:2 CH_2Cl_2:CH_3OH) UV; υ_{max} (solid, cm^{-1}) 2815 (br m), 2574 (w), 2485 (w), 1732 (s), 1631 (m), 1491 (m), 1466 (w), 1422 (m), 1391 (m), 1328 (w), 1303 (w), 1186 (s), 1123 (m); δ_H (300 MHz, D_2O) 8.97 (1H, s, $\mathbf{H}^{2'}$), 8.72 (1H, d, *J*=7.0 Hz, $\mathbf{H}^{4'}$), 8.58 (1H, d, *J*=7.0 Hz, $\mathbf{H}^{6'}$), 8.01 (1H, dd, *J*=7.0 Hz, $\mathbf{H}^{5'}$), 5.36 (2H, s, $\mathbf{CH_2}$); δ_C (75 MHz, D_2O) 171.42 (0), 148.96 (1), 147.60 (1), 146.87 (1), 138.12 (0), 131.07 (1), 64.45 (2); LRMS (ES+) *m/z* 172.0, 174.0 $(M)^+$ (Cl isotope pattern) (45 %), 194.0, 196.0 $(M+Na-H)^+$ (Cl isotope pattern) (100 %); HRMS (ES+) *m/z* 172.0158 $(M)^+$; $C_7H_7NO_2Cl^{35}$ requires 172.0159.

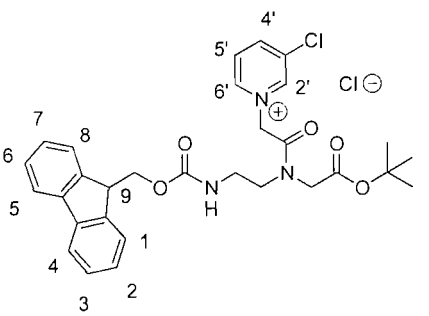
tert-Butyl 2-[[2-(3'-bromo-1'-pyridiniumyl)acetyl](2-[(9*H*-9-fluorenylmethoxy)carbonyl]aminoethyl)amino]acetate chloride (42)



PNA backbone **(22)** (200 mg, 0.46 mmol) was added to a solution of acid **(40)** (153 mg, 0.52 mmol), EDC (120 mg, 0.62 mmol) and DIPEA (70 μL, 0.46 mmol) in anhydrous DMF (10 mL). The reaction mixture was stirred for 1 hour under an atmosphere of argon. The mixture was diluted with CH_2Cl_2 (20 mL) and washed with saturated KCl (50 mL). The organic layer was dried over $MgSO_4$ and concentrated *in vacuo*. The residue was by purified by silica gel column chromatography (CH_2Cl_2 to 2:8 CH_2Cl_2:CH_3OH). Fractions containing the product were concentrated *in vacuo* to yield monomer **(42)** as a brown foam (125 mg, 43 %).

Data: R_f = 0.25 (5:1:1 ethyl acetate:CH_3OH:NH_3) UV; δ_H (400 MHz, $CDCl_3$) (2:1 rotamers) 9.60-9.63 (2H, m, $\mathbf{H}^{6'}$), 9.39-9.42 (1H, m, $\mathbf{H}^{6'}$), 9.12-9.14 (3H, m, $\mathbf{H}^{2'}$), 8.33-8.36 (3H, m, $\mathbf{H}^{4'}$), 7.62-7.70 (12H, m, N**H**, $\mathbf{H}^{5'}$, $\mathbf{H}^{1/8}$, $\mathbf{H}^{4/5}$), 7.52-7.55 (1H, m, $\mathbf{H}^{5'}$), 7.27-7.32 (6H, m, $\mathbf{H}^{3/6}$), 7.18-7.24 (6H, m, $\mathbf{H}^{2/7}$), 6.53 (4H, s, $NC\mathbf{H}_2CO_2(CH_3)_3$), 6.20 (2H, s, $NC\mathbf{H}_2CO_2(CH_3)_3$), 4.48 (4H, s, $NCOC\mathbf{H}_2N$), 4.18-4.26 (9H, m, $C\mathbf{H}_2CH^9$, $C\mathbf{H}^9$), 4.01 (2H, s, $NCOC\mathbf{H}_2N$), 3.58-3.65 (4H, m, $NHCH_2C\mathbf{H}_2N$), 3.48-3.56 (6H, m, $NHC\mathbf{H}_2CH_2N$), 3.32-3.36 (2H, m, $NHC\mathbf{H}_2CH_2N$), 1.42 (18H, s, $C(C\mathbf{H}_3)_3$), 1.37 (9H, s, $C(C\mathbf{H}_3)_3$); δ_C (100 MHz, $CDCl_3$) 169.87 (0), 168.15 (0), 164.67 (0), 148.38 (1), 148.14 (1), 147.62 (1), 145.36 (1), 144.61 (0), 141.61 (0) 136.63 (0), 128.10 (1), 127.67 (1), 127.59 (1), 126.08 (1), 83.10 (0), 67.35 (2), 61.87 (2), 55.32 (2), 49.77 (2), 48.87 (2), 47.36 (1), 29.69 (3), 28.47 (3); LRMS (ES+) *m/z* 594.2, 596.2 $(M)^+$ (Br isotope pattern) (100 %); HRMS (ES+) *m/z* 594.1590 $(M)^+$; $C_{30}H_{33}N_3O_5Br^{79}$ requires 594.1598.

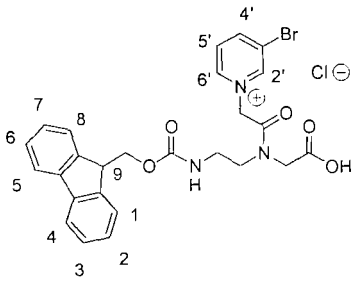
tert-Butyl 2-[[2-(3'-chloro-1'-pyridiniumyl)acetyl](2-[(9_H_-9-fluorenylmethoxy)carbonyl]aminoethyl)amino]acetate chloride (43)



PNA backbone **(22)** (200 mg, 0.46 mmol) was added to a solution of acetic acid **(41)** (130 mg, 0.52 mmol), EDC (120 mg, 0.62 mmol) and DIPEA (70 μL, 0.46 mmol) in anhydrous DMF (10 mL). The reaction mixture was stirred for 1 hour under an atmosphere of argon. The mixture was diluted with CH_2Cl_2 (20 mL) and washed with saturated KCl (50 mL). The organic layer was dried over $MgSO_4$ and concentrated *in vacuo*. The residue was by purified by silica gel column chromatography (CH_2Cl_2 to 2:8 CH_2Cl_2:CH_3OH). Fractions containing the product were concentrated *in vacuo* to yield monomer **(43)** as a brown foam (153 mg, 58 %).

Data: R_f = 0.25 (5:1:1 ethyl acetate:CH_3OH:NH_3) UV; υ_{max} (Film, cm^{-1}) 1712 (s), 1666 (s), 1532 (w), 1493 (w), 1450 (m), 1368 (w), 1248 (s), 1215 (m), 1103 (w), 1033 (w), 1009 (w); δ_H (400 MHz, $CDCl_3$) (1:2 rotamers) 9.551-9.55 (1H, m, $\mathbf{H}^{6'}$), 9.34-9.39 (2H, s, $\mathbf{H}^{6'}$), 9.13-9.15 (1H, m, $\mathbf{H}^{2'}$), 9.08-9.11 (2H, m, $\mathbf{H}^{2'}$), 8.16-8.24 (3H, m, $\mathbf{H}^{4'}$), 7.78-7.80 (1H, m, $\mathbf{H}^{5'}$), 7.62-7.70 (15H, m, N**H**, $\mathbf{H}^{5'}$, $\mathbf{H}^{1/8}$, $\mathbf{H}^{4/5}$), 7.28-7.33 (6H, m, $\mathbf{H}^{3/6}$), 7.21-7.28 (6H, m, $\mathbf{H}^{2/7}$), 6.76-6.69 (2H, m, N**H**), 6.57 (2H, s, $NC\mathbf{H}_2CO_2(CH_3)_3$), 6.20 (4H, s, $NC\mathbf{H}_2CO_2(CH_3)_3$), 4.46 (2H, s, $NCOC\mathbf{H}_2N$), 4.08-4.24 (9H, m, $C\mathbf{H}_2CH^9$, $C\mathbf{H}^9$), 4.00 (4H, s, $NCOC\mathbf{H}_2N$), 3.58-3.65 (2H, m, $NHCH_2C\mathbf{H}_2N$), 3.51-3.58 (6H, m, $NHC\mathbf{H}_2CH_2N$, $NHCH_2C\mathbf{H}_2N$), 3.30-3.38 (4H, m, $NHC\mathbf{H}_2CH_2N$), 1.42 (9H, s, $(C\mathbf{H}_3)_3$), 1.36 (18H, s, $(C\mathbf{H}_3)_3$); δ_C (100 MHz, $CDCl_3$) 169.52 (0), 168.11 (0), 164.82 (0), 146.27 (1), 145.98 (1), 145.66 (1), 145.56 (1), 145.53 (1), 144.37 (0), 141.61 (0), 135.49 (0), 128.09 (1), 127.58 (1), 120.27 (1), 84.06 (0), 67.25 (2), 61.93 (2), 54.30 (2), 49.75 (2), 48.91 (2), 47.61 (1), 31.29 (3), 28.46 (3); LRMS (ES+) *m/z* 550.2, 552.2 $(M)^+$ (Cl isotope pattern) (100 %); HRMS (ES+) *m/z* 550.2099 $(M)^+$; $C_{30}H_{33}N_3O_5Cl^{35}$ requires 550.2104.

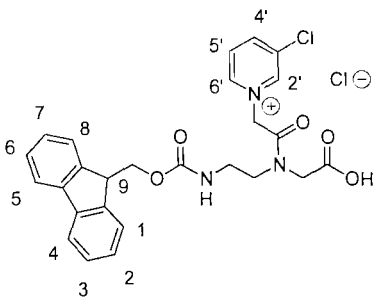
2-[[2-(3'-Bromo-1'-pyridiniumyl)acetyl](2-[(9*H*-9-fluorenylmethoxy)carbonyl]aminoethyl)amino]acetic acid chloride (44)



Monomer **(42)** (95 mg, 0.15 mmol) was added to a solution of TFA:CH_2Cl_2 (1:1) (10 mL). The reaction mixture was stirred for 1 hour at room temperature. The mixture was concentrated *in vacuo* and the residue was triturated with Et_2O. The resulting brown solid was isolated by filtration and washed with Et_2O to yield monomer **(44)** as a brown solid (90 mg, 96 %).

Data: R_f = 0.07 (5:1:1 ethyl acetate:CH_3OH:NH_3) UV; δ_H (400 MHz, CD_3OD) (approx.1:1 rotamers) 9.11 (1H, s, $\mathbf{H}^{2'}$), 9.07 (1H, s, $\mathbf{H}^{2'}$), 8.67-8.79 (4H, m, $\mathbf{H}^{4'}$, $\mathbf{H}^{6'}$), 7.92-7.99 (1H, m, $\mathbf{H}^{5'}$), 7.84-7.88 (1H, m, $\mathbf{H}^{5'}$), 7.68-7.72 (4H, m, $\mathbf{H}^{4/5}$), 7.49-7.57 (4H, m, $\mathbf{H}^{1/8}$), 7.24-7.31 (4H, m, $\mathbf{H}^{3/6}$), 7.15-7.23 (4H, m, $\mathbf{H}^{2/7}$), 5.67 (2H, s, $NC\mathbf{H}_2CO_2H$), 5.47 (2H, s, $NC\mathbf{H}_2CO_2H$), 4.38-4.03 (10H, m, $NCOC\mathbf{H}_2N$, $C\mathbf{H}_2CH^9$, $C\mathbf{H}^9$), 3.34-3.51 (4H, m, $NHCH_2C\mathbf{H}_2N$), 3.31-3.42 (2H, m, $NHC\mathbf{H}_2CH_2N$), 3.18-3.24 (2H, m, $NHC\mathbf{H}_2CH_2N$); δ_C (100 MHz, CD_3OD) 181.25 (0), 172.76 (0), 168.57 (0), 150.85 (0), 150.69 (1), 149.54 (1), 147.08 (1) 145.62 (0), 143.02 (0), 129.90 (1), 129.25 (1), 128.58 (1), 126.58 (1), 121.40 (1), 70.58 (2), 68.17 (2), 67.29 (2), 66.53 (2), 63.45 (2), 40.36 (1); LRMS (ES+) *m/z* 538.0, 540.0 $(M)^+$ (Br isotope pattern) (100 %); HRMS (ES+) *m/z* 538.0965 $(M)^+$; $C_{26}H_{25}N_3O_5Br^{79}$ requires 538.0973.

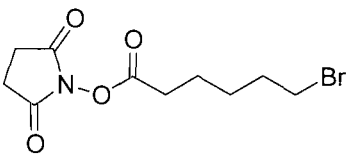
2-[[2-(3'-Chloro-1'-pyridiniumyl)acetyl](2-[(9*H*-9-fluorenylmethoxy)carbonyl]aminoethyl)amino]acetic acid chloride (45)



Monomer **(43)** (120 mg, 0.20 mmol) was added to a solution of TFA:CH_2Cl_2 (1:1) (10 mL). The reaction mixture was stirred for 1 hour at room temperature. The mixture was concentrated *in vacuo* and the residue was triturated with Et_2O. The resulting brown solid was isolated by filtration and washed with Et_2O to yield monomer **(45)** as a brown solid (97 mg, 92 %).

Data: R_f = 0.06 (5:1:1 ethyl acetate:CH_3OH:NH_3) UV; δ_H (400 MHz, CD_3OD) (approx.1:1 rotamers) 9.04 (1H, s, $\mathbf{H}^{2'}$), 9.00 (1H, s, $\mathbf{H}^{2'}$), 8.67-6.75 (4H, m, $\mathbf{H}^{4'}$, $\mathbf{H}^{6'}$), 8.01-8.08 (1H, m, $\mathbf{H}^{5'}$), 7.94-7.98 (1H, m, $\mathbf{H}^{5'}$), 7.66-7.72 (4H, m, $\mathbf{H}^{4/5}$), 7.52-7.57 (4H, m, $\mathbf{H}^{1/8}$), 7.24-7.32 (4H, m, $\mathbf{H}^{3/6}$), 7.17-7.24 (4H, m, $\mathbf{H}^{2/7}$), 5.69 (2H, s, $NC\mathbf{H}_2CO_2H$), 5.49 (2H, s, $NC\mathbf{H}_2CO_2H$), 4.37-4.03 (10H, m, $NCOC\mathbf{H}_2N$, $C\mathbf{H}_2CH^9$, $C\mathbf{H}^9$), 3.42-3.49 (4H, m, $NHCH_2C\mathbf{H}_2N$), 3.30-3.38 (2H, m, $NHC\mathbf{H}_2CH_2N$), 3.18-3.24 (2H, m, $NHC\mathbf{H}_2CH_2N$); δ_C (100 MHz, CD_3OD) 210.90 (0), 169.23 (0), 162.16 (0), 159.35 (0), 151.262 (1), 149.39 (0), 146.51 (1), 137.83 (0), 134.80 (1), 132.55 (1), 130.60 (1), 128.68 (1), 123.56 (1), 113.78 (1), 69.98 (2), 69.03 (2), 67.97 (2), 60.74 (2), 55.21 (2), 37.85 (1); LRMS (ES+) *m/z* 494.0, 496.0 $(M)^+$ (Cl isotope pattern) (100 %); 516.1, 518.1 $(M^+\text{-}H\text{+}Na)^+$ (Cl isotope pattern) (80 %); HRMS (ES+) *m/z* 494.1481 $(M)^+$; $C_{26}H_{25}N_3O_5Cl^{35}$ requires 494.1477.

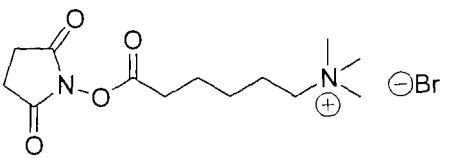
2,5-Dioxotetrahydro-1*H*-1-pyrrolyl 6-bromohexanoate (47)



Following a procedure described by Bartlet-Jones *et al.*[145] 6-bromohexanoyl chloride **(46)** (6.50 mL, 42.47 mmol) was added to a solution of *N*-hydroxysuccinimide (5.38 g, 46.72 mmol) in anhydrous acetonitrile (30 mL). The reaction mixture was refluxed for 12 hours under an atmosphere of argon. The reaction mixture was then concentrated *in vacuo* and recrystallised from ethanol to yield ester **(47)** as a colourless solid (9.46 g, 76 %).

Data: R_f = 0.58 (9:1 CH_2Cl_2:CH_3OH); UV and phosphomolybdic acid; υ_{max} (Film, cm^{-1}) 2944 (w), 2864 (w), 1812 (w), 1783 (m), 1735 (s), 1461 (w), 1429 (w), 1365 (w), 1205 (m), 1068 (m); δ_H (400 MHz, $CDCl_3$) 3.45 (2H, t, *J*=7.0 Hz, CH_2Br), 2.86 (4H, s, $COCH_2CH_2CO$), 2.67 (2H, t, *J*=7.5 Hz, $COCH_2$), 1.89-1.95 (2H, dt, *J*=7.0, 7.0 Hz, CH_2CH_2Br), 1.78-1.87 (2H, dt, *J*=7.0, 7.5 Hz, $COCH_2CH_2$), 1.62-1.64 (2H, m, $CH_2CH_2CH_2Br$); δ_C (100 MHz, $CDCl_3$) 170.54 (0), 169.78 (0), 34.56 (2), 33.59 (2), 32.17 (2), 28.70 (2), 27.00 (2), 25.30 (2); LRMS (ES+) *m/z* 314.0, 316.0 $(M+Na)^+$ (Br isotope pattern 100 %); HRMS (ES+) *m/z* 346.0262 $(M+Na+CH_3OH)^+$; $C_{11}H_{18}NO_5Br^{79}Na$ requires 346.0261.

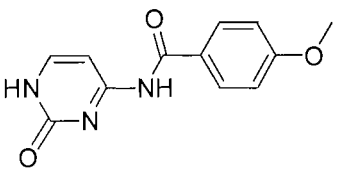
6-[(2,5-Dioxotetrahydro-1*H*-1-pyrrolyl)oxy]-6-oxohexyl(trimethyl)ammonium bromide (4)



Following a procedure described by Bartlet-Jones *et al.*[145] Trimethylamine (1.45 mL, 15.40 mmol) in anhydrous tetrahydrofuran (THF) (3 mL) was added to a solution of ester **(47)** (1.50 g, 5.13 mmol) in anhydrous THF (3 mL). The reaction mixture was stirred for 12 hours under an atmosphere of argon. The resulting colourless solid was filtered and recrystallised in hot ethanol to yield ester **(4)** as a colourless solid (1.35 g, 97 %).

Data: R_f = 0.14 (8:2 CH_2Cl_2:CH_3OH); UV and ninhydrin; υ_{max} (Film, cm^{-1}) 2931 (br w), 1812 (m), 1776 (m), 1720 (s), 1494 (w), 1481 (w), 1469 (w), 1420 (w), 1371 (m), 1211 (s), 1056 (s), 1003 (w); δ_H (400 MHz, CD_3OD) 3.15-3.19 (2H, m, CH_2N), 3.00 (9H, s, $N(CH_3)_3$), 2.64 (4H, s, $COCH_2CH_2CO$), 2.51 (2H, t, *J*=7.5 Hz, $COCH_2CH_2$), 1.65-1.76 (4H, m, $CH_2CH_2CH_2CH_2N$), 1.28-1.39 (2H, d t, *J*=7.5, 8.0 Hz, $COCH_2CH_2CH_2$); δ_C (100 MHz, CD_3OD) 174.46 (0), 172.72 (0), 70.17 (2), 56.16 (3), 33.78 (2), 29.08 (2), 28.94 (2), 27.58 (2), 25.97 (2); LRMS (ES+) *m/z* 271.1 $(M)^+$ (100 %); HRMS (ES+) *m/z* 271.1652 $(M)^+$; $C_{13}H_{23}N_2O_4$ requires 271.1652.

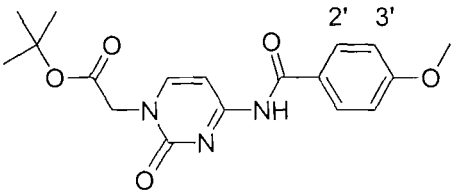
_N_-1-(2-Oxo-1,2-dihydro-4-pyrimidinyl)-4-methoxybenzamide (49)



Following the procedure described by Will *et al.*[164] 4-Anisoyl chloride (32.55 g, 208 mmol) in anhydrous pyridine (100 mL) was added to a rapidly stirred suspension of cytosine **(48)** (20.8 g, 181 mmol) in anhydrous pyridine (50 mL) under an atmosphere of argon. After 50 minutes, anhydrous DMF (100 mL) was added to the white suspension and the reaction was stirred for a further 30 minutes. The solvent was removed *in vacuo* and the residue suspended in DMF/Et_2O 1:1 (1 L). The colourless solid was collected by filtration, washed with water and Et_2O and to yield compound **(49)** as a colourless solid, which was dried *in vacuo* over P_2O_5 (39.26 g, 89 %). Due to the poor solubility of ester **(49)** in organic solvents little characterisation was performed. The data obtained was found to be consistent Will *et al.*[164]

Data: R_f = 0.75 (9:1 CH_2Cl_2:CH_3OH) UV and ninhydrin; MP = 314-316 °C; υ_{max} (solid, cm^{-1}) 3244 (w), 3148 (w), 3071 (w), 2844 (w), 1687 (s), 1618 (s), 1606 (m), 1489 (s), 1449 (s), 1425 (s), 1405 (m), 1325 (w), 1302 (m), 1256 (s), 1241 (s), 1214 (m), 1185 (s), 1119 (m), 1084 (m), 1020 (m); LRMS (ES+) *m/z* 246.1 $(M+H)^+$ (100%).

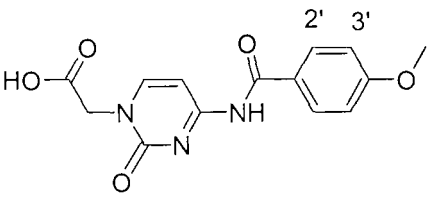
tert-Butyl 2-4-[(4'-methoxybenzoyl)amino]-2-oxo-1,2-dihydro-1-pyrimidinylacetate (50)



Using a method adapted from the procedure described by Thomson *et al.*[160] Anhydrous K_2CO_3 (16.17 g, 117 mmol) and anhydrous Cs_2CO_3 (3.18 g, 1 mmol) were added to a suspension of ester **(49)** (28.45 g, 116 mmol) in anhydrous DMF (450 mL). After 10 minutes *tert*-butyl bromoacetate (17 mL, 116 mmol) was added drop wise and the mixture was stirred for 48 hours. The resulting suspension was filtered and the filtrate concentrated *in vacuo* to yield an orange solid. Addition of ethyl acetate (200 mL) yielded a white solid, which was isolated by filtration and washed with ethyl acetate (100 mL). The filtrate was concentrated *in vacuo* to approximately 100 mL of ethyl acetate and stored overnight at –20 °C to yield ester **(50)** as a colourless solid which was dried overnight *in vacuo* over P_2O_5 (28.93 g, 69 %).

Data: R_f = 0.47 (1:1 CH_2Cl_2:ethyl acetate) UV and ninhydrin; MP = 194-196 °C; υ_{max} (solid, cm^{-1}) 2980 (w), 2938 (w), 2841 (w), 2362 (w), 1741 (m), 1664 (s), 1628 (m), 1605 (m), 1576 (m), 1556 (m), 1484 (s), 1440 (w), 1407 (w), 1361 (s), 1303 (m), 1241 (s), 1173 (m), 1156 (s), 1105 (w), 1081 (w), 1029 (w); δ_H (400 MHz, d_6-DMSO) 11.20 (1H, s, N**H**), 8.25 (1H, d, *J*=7.6 Hz, **H**6), 8.15 (2H, d, *J*=8.7 Hz, $\mathbf{H}^{2'}$), 7.40 (1H, d, *J*=7.6 Hz, **H**5), 7.15 (2H, d, *J*=8.7 Hz, $\mathbf{H}^{3'}$), 4.65 (2H, s, NC$\mathbf{H}_2$CO), 3.95 (3H, s, ArOC$\mathbf{H}_3$), 1.55 (9H, s, C(C$\mathbf{H}_3$)$_3$); δ_C (100 MHz, d_6 -DMSO) 167.43 (0), 166.87 (0), 164.19 (0), 163.30 (0), 155.70 (0), 150.96 (1), 131.08 (1), 125.59 (0), 114.20 (1), 96.26 (1), 82.21 (0), 56.01 (3), 51.65 (2), 28.25 (3); LRMS (ES+) *m/z* 360.2 $(M+H)^+$ (20%), 382.2 $(M+Na)^+$ (100%), 719.4 $(2M+H)^+$ (20%); HRMS (ES+) *m/z* 382.1368 $(M+Na)^+$; $C_{18}H_{21}N_3O_5Na$ requires 382.1373.

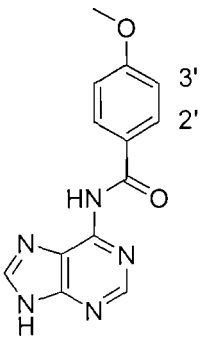
2-4-[(4'-Methoxybenzoyl)amino]-2-oxo-1,2-dihydro-1-pyrimidinylacetic acid (51)



Using a method adapted from the procedure described by Thomson *et al.*[160] 4 M HCl (34 mL) in 1,4-dioxane (100 mL) was added to a rapidly stirred suspension of ester **(50)** (27.27 g, 76 mmol) in CH_2Cl_2 (100 mL). The thick white suspension was stirred overnight. The suspension was filtered, washed with Et_2O and dried over P_2O_5 to yield acid **(51)** (22.51 g, 98 %). The data obtained was found to be consistent with Thomson *et al.*[160]

Data: R_f = 0.03 (9:1 CH_2Cl_2:CH_3OH) UV and ninhydrin; MP = 238-240 °C; υ_{max} (solid, cm^{-1}) 3299 (w), 3076 (w), 2913 (w), 2495 (w), 1714 (m), 1664 (m), 1630 (m), 1605 (m), 1597 (s), 1573 (m), 1543 (m), 1474 (w), 1459 (w), 1442 (w), 1405 (w), 1371 (w), 1349 (w), 1322 (m), 1268 (s), 1251 (s), 1236 (m), 1193 (m), 1182 (s), 1162 (m), 1116 (m) 1069 (m); δ_H (300 MHz, d_6-DMSO) 8.40 (1H, d, *J*=8.0 Hz, **H**6), 8.10 (2H, d, *J*=9.0 Hz, **H**$^{2'}$), 7.45 (1H, d, *J*=8.0 Hz, **H**5), 7.05 (2H, d, *J*=9.0 Hz, **H**$^{3'}$), 5.40 (1H, br s, N**H**), 4.65 (2H, s, NC$\mathbf{H}_2$CO), 3.80 (3H, s, ArOC$\mathbf{H}_3$); δ_C (75 MHz, d_6 -DMSO) 168.79 (0), 166.11 (0), 163.50 (0), 161.71 (0), 153.93 (1), 151.39 (0), 131.04 (1), 124.23 (0), 114.11 (1), 95.52 (1), 55.74 (3), 50.81 (2); LRMS (ES+) *m/z* 304.1 $(M+H)^+$ (100 %); (ES-) *m/z* 302.2 $(M-H)^-$ (10 %).

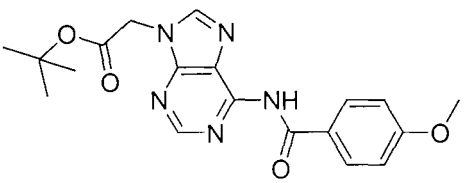
_N_-1-(9_H_-6Purinyl)-4'-methoxybenzamide (53)



Following the procedure described by Will *et al.*[164] 4-Anisoyl chloride (21.50 g, 159 mmol) in anhydrous pyridine (100 mL) was added to a rapidly stirred suspension of adenine **(52)** (32.57 g, 191 mmol) in anhydrous pyridine (50 mL) at 0 °C under an atmosphere of argon. The mixture was heated to 100 °C for 4 hours the solvent was removed *in vacuo* (azeotroped with toluene). The orange solid was stirred in refluxing propan-2-ol (250 mL) for 1 hour. The white suspension was collected by filtration and dried *in vacuo* over P_2O_5 to yield compound **(53)** (36.02 g, 84 %). The data obtained was found to be consistent with Will *et al.*[164]

Data: R_f = 0.53 (5:1:1 ethyl acetate:CH_3OH:NH_3) UV and ninhydrin; MP = 216-218 °C; υ_{max} (solid, cm^{-1}) 3297 (w), 1699 (w), 1661 (m), 1605 (m), 1556 (w), 1525 (w), 1498 (m), 1459 (s), 1436 (m), 1401 (m), 1376 (m), 1328 (m), 1318 (m), 1280 (w), 1251 (s), 1217 (m), 1177 (s), 1110 (m), 1068 (w), 1031 (m); δ_H (300 MHz, d_6-DMSO) 11.50 (1H, br s, N**H**CO), 9.00 (1H, br s, N**H**), 8.70 (1H, s, **H**2), 8.55 (1H, s, **H**8), 8.15 (2H, d, *J*=8.9 Hz, $\mathbf{H}^{2'}$), 7.10 (2H, d, *J*=8.9 Hz, $\mathbf{H}^{3'}$), 3.85 (3H, s, ArOC$\mathbf{H}_3$); δ_C (75 MHz, d_6-DMSO) 165.88 (0), 162.89 (0), 151.19 (1), 145.76 (1), 145.39 (0), 143.15 (0), 130.84 (1), 124.83 (0), 115.35 (0), 113.85 (1), 55.63 (3); LRMS (ES+) *m/z* 270.1 $(M+H)^+$ (100 %), 539.2 $(2M+H)^+$ (10 %).

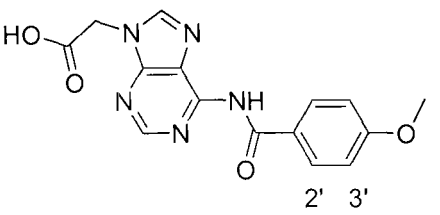
tert-Butyl 2-[6-[(4'-methoxybenzoyl)amino]-9*H*-9-purinyl]acetate (54)



Following a method adapted from the procedure described by Thomson *et al.*[160] Anhydrous K_2CO_3 (18.2 g, 132 mmol) and anhydrous Cs_2CO_3 (4.3 g, 1.3 mmol) were added to a suspension of compound **(53)** (35.66 g, 132 mmol) in anhydrous DMF (250 mL). After 10 min *tert*-butyl bromoacetate (21.51 mL, 145 mmol) was added drop wise and the mixture was stirred for 48 hours. The resulting suspension was filtered. The filtrate was concentrated *in vacuo* into approximately 100 mL of ethyl acetate yielding a pale brown solid, which was isolated by filtration and washed with ethyl acetate, water then Et_2O. The solid was dissolved in the minimum quantity of ethyl acetate and boiled with activated charcoal. The mixture was filtered and the filtrate stored overnight at 5 °C. The resulting colourless solid was isolated and washed with cold Et_2O and dried *in vacuo* over P_2O_5 to give ester **(54)** as a colourless solid (6.03 g, 12 %).

Data: R_f = 0.80 (5:1:1 ethyl acetate:CH_3OH:NH_3) UV and ninhydrin; MP = 137-139 °C; υ_{max} (solid, cm^{-1}) 1741 (m), 1688 (m), 1606 (m), 1582 (s), 1516 (w), 14814 (w), 1466 (w), 1444 (m), 1405 (w), 1374 (w), 1365 (w), 1327 (w), 1288 (w), 1240 (s), 1188 (m), 1165 (s), 1157 (m), 1122 (w), 1093 (m), 1030 (m); δ_H (300 MHz, d_6-DMSO) 11.05 (1H, s, N**H**), 8.70 (1H, s, **H**2), 8.40 (1H, s, **H**8), 8.10 (2H, d, *J*=9.0 Hz, $\mathbf{H}^{2'}$), 7.05 (2H, d, *J*=9.0 Hz, $\mathbf{H}^{3'}$), 5.10 (2H, s, $NC\mathbf{H}_2CO$), 3.90 (3H, s, $ArOC\mathbf{H}_3$), 1.45 (9H, s, $C(C\mathbf{H}_3)_3$); δ_C (75 MHz, d_6-DMSO) 166.85 (0), 164.98 (0), 162.65 (0), 152.60 (0), 151.73 (1), 150.50 (0), 145.01 (1), 130.67 (1), 125.55 (0), 124.97 (0), 113.78 (1), 82.40 (0), 55.57 (3), 44.89 (2), 27.73 (3); LRMS (ES+) *m/z* 384.1 $(M+H)^+$ (100 %), 406.1 $(M+Na)^+$ (50 %); HRMS (ES+) *m/z* 384.1655 $(M+H)^+$; $C_{19}H_{22}N_5O_4$ requires 384.1666.

2-6-[(4'-Methoxybenzoyl)amino]-9*H*-9-purinylacetic acid (55)



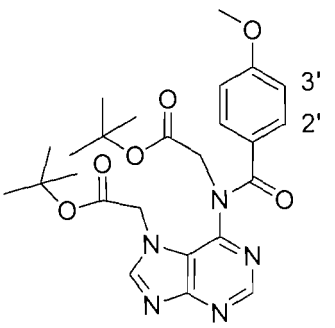
Using a method adapted from the procedure described by Thomson *et al.*[160] TFA (51 mL) was added drop wise to a solution of compound **(54)** (5.08 g, 13 mmol) and triethylsilane (20 mL) in anhydrous CH_2Cl_2 (150 mL) and cooled to 0 °C. The reaction was allowed to warm to room temperature and was rapidly stirred for 7 hours. The mixture was evaporated to dryness (azeotroped with chloroform) and was dried *in vacuo* over KOH yielding acid **(55)** as a pale yellow solid (4.30 g, 99 %). The data obtained was found to be consistent with Thomson *et al.*[160]

Data: R_f = 0.70 (10:2:1 CH_2Cl_2:CH_3OH:glacial acetic acid) UV and ninhydrin; MP = 224-227 °C; υ_{max} (film, cm^{-1}) 1735 (w), 1686 (m), 1653 (m), 1650 (m), 1604 (m), 1551 (w), 1507 (w), 1460 (w), 1420 (w), 1356 (w), 1313 (w), 1257 (s), 1176 (s), 1148 (m), 1102 (w), 1181 (w), 1026 (w); δ_H (300 MHz, d_6 - DMSO) 11.20 (1H, br s, CO_2**H**), 8.75 (1H, s, **H**2), 8.65 (1H, s, **H**8), 8.05 (2H, d, *J*=9.0 Hz, **H**$^{2'}$), 7.10 (2H, d, *J*=9.0 Hz, **H**$^{3'}$), 5.15 (2H, s, NC**H**$_2$CO), 3.90 (3H, s, ArOC**H**$_3$); δ_C (75 MHz, d_6-DMSO) 166.99 (0), 165.42 (0), 162.81 (0), 152.29 (0), 151.74 (1), 149.93 (0), 145.06 (1), 130.80 (1), 125.26 (0), 123.16 (0), 113.82 (1), 55.58 (3), 44.59 (2); LRMS (ES+) *m/z* 328.3 $(M+H)^+$ (30 %), 350.3 $(M+Na)^+$ (100 %); (ES-) *m/z* 326.3 $(M+H)^-$ (100 %).

tert-Butyl 2-[6-[[2-(_tert_-butoxy)-2-oxoethyl](4'-methoxybenzoyl)amino]-7_H_-7-purinyl] acetate (56) and _tert_-butyl 2-6-[[2-(_tert_-butoxy)-2-oxoethyl](4'-methoxybenzoyl)amino]-9_H_-9-purinylacetate (57)

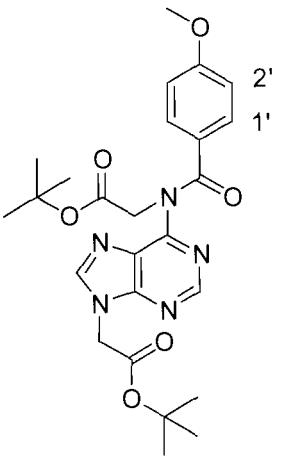
Anhydrous K_2CO_3 (12.02 g, 87 mmol) and anhydrous Cs_2CO_3 (2.89 g, 9 mmol) were added to a suspension of compound **(53)** (23.44 g, 87 mmol) in anhydrous DMF (500 mL). After 10 minutes *tert*-butyl bromoacetate (14 mL, 96 mmol) was added drop wise and the mixture was stirred overnight then heated to 60 °C for 4 hours. The resulting suspension was filtered and the filtrate concentrated *in vacuo* to yield an orange solid. The residue was dissolved in ethyl acetate and extracted with water (2 x 200 mL). The organic layer was dried over Na_2SO_4 and filtered. The filtrate was boiled with activated charcoal, filtered and stored overnight at 5 °C yielding ester **(56)** as a colourless solid (10.61 g, 24 %). Concentration of the filtrate *in vacuo* yielded ester **(57)** as a white solid (11.32 g, 26 %).

Data for *tert*-butyl 2-[6-[[2-(*tert*-butoxy)-2-oxoethyl](4'-methoxybenzoyl)amino]-7*H*-7-purinyl] acetate (56)



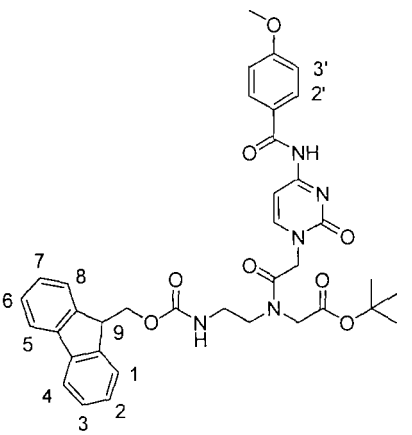
Data: R_f = 0.78 (5:1:1 ethyl acetate:CH_3OH:NH_3) UV and ninhydrin; MP = 160-162 °C; υ_{max} (solid, cm^{-1}) 2362 (w), 1745 (m), 1637 (m), 1600 (m), 1573 (w), 1547 (w), 1508 (m), 1478 (w), 1458 (w), 1436 (w), 1415 (w), 1390 (w), 1366 (m), 1311 (m), 1272 (m), 1235 (s), 1151 (s), 1112 (m), 1102 (m), 1030 (m), 1015 (m), 1005 (m); δ_H (300 MHz, d_6-DMSO) 8.35 (1H, s, **H**2), 7.90 (1H, s, **H**8), 7.80 (2H, d, *J*=9.0 Hz, $\mathbf{H}^{2'}$), 6.95 (2H, d, *J*=9.0 Hz, $\mathbf{H}^{3'}$), 5.00 (2H, s, $NC\mathbf{H}_2CO$), 4.85 (2H, s, $N9C\mathbf{H}_2CO$), 3.80 (3H, s, $ArOC\mathbf{H}_3$), 1.45 (9H, s, $CO_2(C\mathbf{H}_3)_3$), 1.40 (9H, s, $N9CH_2CO_2(C\mathbf{H}_3)_3$); δ_C (75 MHz, d_6-DMSO) 178.12 (0), 166.99 (0), 166.59 (0), 162.11 (0), 148.26 (0), 147.21 (1), 144.86 (0), 142.72 (1), 130.99 (1), 128.18 (0), 113.28 (1), 112.14 (0), 82.33 (0), 81.94 (0), 55.29 (3), 48.53 (2), 48.10 (2), 27.64 (3), 27.58 (3); LRMS (ES+) *m/z* 498.4 $(M+H)^+$ (100 %), 520.4 $(M+Na)^+$ (60 %); HRMS (ES+) *m/z* 498.2338 $(M+H)^+$; $C_{25}H_{32}N_5O_6$ requires 498.2347.

Data for *tert*-butyl 2-6-[[2-(*tert*-butoxy)-2-oxoethyl](4'-methoxybenzoyl)amino]-9*H*-9-purinylacetate (57)



Data: R_f = 0.78 (5:1:1 ethyl acetate:CH_3OH:NH_3) UV and ninhydrin; MP = 184-186 °C; υ_{max} (solid, cm^{-1}) 2362 (w), 1745 (m), 1637 (m), 1596 (m), 1571 (m), 1503 (m), 1470 (m), 1457 (m), 1413 (w), 1401 (w), 1385 (m), 1370 (m), 1338 (w), 1309 (m), 1267 (m), 1235 (s), 1158 (s), 1113 (m), 1059 (m), 1031 (m), 1014 (m); δ_H (300 MHz, d_6-DMSO) 8.20 (1H, s, **H**2), 8.05 (1H, s, **H**8), 8.85 (2H, d, *J*=8.8 Hz, $\mathbf{H}^{2'}$), 7.95 (2H, d, *J*=8.8 Hz, $\mathbf{H}^{3'}$), 5.20 (2H, s, NC$\mathbf{H}_2$CO), 4.85 (2H, s, N9C$\mathbf{H}_2$CO), 3.80 (3H, s, ArOC$\mathbf{H}_3$), 1.40 (9H, s, N9$CH_2CO_2(C\mathbf{H}_3)_3$), 1.35 (9H, s, $CO_2(C\mathbf{H}_3)_3$); δ_C (75 MHz, d_6-DMSO) 178.12 (0), 166.99 (0), 166.59 (0), 162.11 (0), 148.26 (0), 147.21 (1), 144.86 (0), 142.72 (1), 130.99 (1), 128.18 (0), 113.28 (1), 112.14 (0), 82.33 (0), 81.94 (0), 55.29 (3), 48.53 (2), 48.10 (2), 27.64 (3), 27.58 (3); LRMS (ES+) *m/z* 498.2 $(M+H)^+$ (100%); HRMS (ES+) *m/z* 498.2341 $(M+H)^+$; $C_{25}H_{32}N_5O_6$ requires 498.2347.

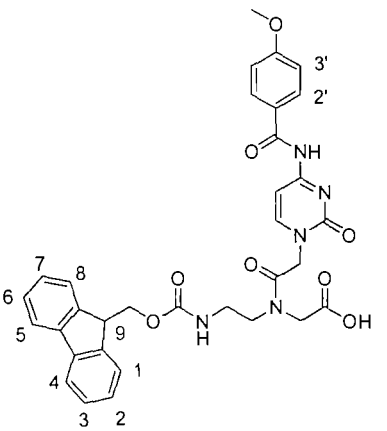
tert-Butyl 2-[(2-[(9*H*-9-fluorenylmethoxy)carbonyl]aminoethyl)(2-4-[(4'-methoxybenzoyl)amino]-2-oxo-1,2-dihydro-1-pyrimidinylacetyl)amino]acetate (58)



Following the procedure described by Thompson *et al.*[160] PNA backbone **(22)** (500 mg, 1.26 mmol) was added to a solution of acid **(51)** (400 mg, 1.32 mmol), EDC (600 mg, 3.14 mmol) and DIPEA (215 μL, 1.26 mmol) in anhydrous DMF (15 mL). The reaction mixture was stirred for 3 hours under an atmosphere of argon. The mixture was diluted with CH_2Cl_2 (50 mL) and washed with saturated KCl (50 mL). The organic layer was dried over $MgSO_4$ and concentrated *in vacuo*. The residue was purified by silica gel chromatography (CH_2Cl_2 to 7:3 CH_2Cl_2:CH_3OH). Fractions containing the product were concentrated *in vacuo* to yield monomer **(58)** as a colourless solid (120 mg, 13 %).

Data: R_f = 0.25 (99:1 ethyl acetate:NH_3) UV; δ_H (400 MHz, d_6-DMSO) (2:1 rotamers) 11.32 (3H, br t, N**H**), 808-8.17 (3H, m, **H**6), 8.02-8.16 (6H, m, $\mathbf{H}^{4/5}$), 7.96-8.00 (2H, m, N**H**), 7.83-7.96 (7H, m, $\mathbf{H}^{2'}$, N**H**), 7.68-7.76 (6H, m, $\mathbf{H}^{1/8}$), 7.40-7.50 (6H, m, $\mathbf{H}^{3/6}$), 7.32-7.39 (9H, m, $\mathbf{H}^{2/7}$, **H**5), 7.01-7.12 (6H, m, $\mathbf{H}^{3'}$), 4.90 (4H, s, $NC\mathbf{H}_2CO_2(CH_3)_3$), 4.68-4.72 (6H, m, $NCOC\mathbf{H}_2N$, $NC\mathbf{H}_2CO_2(CH_3)_3$), 4.59 (2H, s, $NCOC\mathbf{H}_2N$), 4.38-4.45 (4H, m, $C\mathbf{H}_2CH^9$), 4.32-4.37 (2H, m, $C\mathbf{H}_2CH^9$), 4.23-4.31 (3H, m, $C\mathbf{H}^9$), 4.01 (6H, s, $OC\mathbf{H}_3$), 3.89 (3H, s, $OC\mathbf{H}_3$), 3.50-3.25 (12H, m, $NHC\mathbf{H}_2C\mathbf{H}_2N$), 1.47 (27H, 2×s, $C(C\mathbf{H}_3)_3$); LRMS (ES+) *m/z* 682.3 $(M+H)^+$ (100 %), 704.3 $(M+Na)^+$ (70 %); HRMS (ES+) *m/z* 682.2879 $(M+H)^+$; $C_{37}H_{40}N_5O_8$ requires 682.2871.

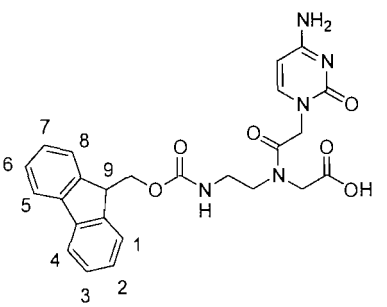
2-[(2-[(9*H*-9-Fluorenylmethoxy)carbonyl]aminoethyl)(2-4-[(4'-methoxybenzoyl)amino]-2-oxo-1,2-dihydro-1-pyrimidinylacetyl)amino]acetic acid (59)



Monomer **(58)** (67 mg, 0.09 mmol) was added to a solution of TFA:CH_2Cl_2 (1:1) (10 mL). The reaction mixture was stirred for 1 hour at room temperature under an atmosphere of argon. The mixture was concentrated *in vacuo* and the residue was triturated with Et_2O. The resulting white solid was isolated by filtration and washed with Et_2O to yield monomer **(59)** as a colourless solid (62 mg, 99 %).

Data: R_f = 0.40 (85:15 CH_2Cl_2:CH_3OH) UV; υ_{max} (solid, cm^{-1}) 1260 (w), 1660 (m), 1603 (m), 1562 (w), 1484 (m), 1450 (w), 1415 (w), 1364 (m), 1307 (w), 1244 (s), 1173 (s), 1140 (m), 1106 (w), 1024 (w); δ_H (300 MHz, d_6-DMSO) (2:1 rotamers) 8.11-8.19 (9H, m, **H**6, $\mathbf{H}^{2'}$), 7.85-8.04 (8H, m, N**H**, $\mathbf{H}^{4/5}$), 7.69-7.78 (6H, m, $\mathbf{H}^{1/8}$), 7.48-7.55 (6H, m, $\mathbf{H}^{3/6}$), 7.30-7.48 (7H, m, $H^{2/7}$, N**H**); 7.08-7.14 (9H, m, H5, $H^{3'}$), 4.97 (4H, s, NC$\mathbf{H}_2$$CO_2H$), 4.94 (2H, s, NC$\mathbf{H}_2$$CO_2H$), 4.69-4.35 (9H, m, C$\mathbf{H}_2$$CH^{9,}$ C$\mathbf{H}^{9}$), 4.12 (4H, s, NC$\mathbf{H}_2$CON), 3.81 (2H, s, NC$\mathbf{H}_2$CON), 3.52-3.58 (4H, m, $NHCH_2C\mathbf{H}_2N$), 3.44-3.51 (2H, m, $NHCH_2C\mathbf{H}_2N$), 3.38-3.44 (4H, m, $NHC\mathbf{H}_2CH_2N$), 3.20-3.27 (2H, m, $NHC\mathbf{H}_2CH_2N$), 2.61 (9H, br s, OC$\mathbf{H}_3$); LRMS (ES+) *m/z* 626.2 $(M+H)^+$ (100 %), 648.2 $(M+Na)^+$ (50 %); HRMS (ES+) *m/z* 626.2248 $(M+H)^+$; $C_{33}H_{32}N_5O_8$ requires 626.2245.

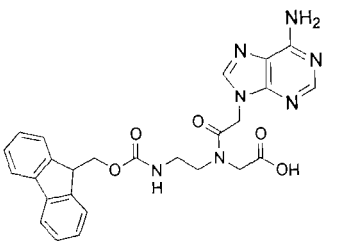
2-[[2-(4-Amino-2-oxo-1,2-dihydro-1-pyrimidinyl)acetyl](2-[(9*H*-9-fluorenylmethoxy)carbonyl]aminoethyl)amino]acetic acid (61)



Monomer **(60)** (480 mg, 0.68 mmol) was added to a solution of TFA:*m*-cresol (4:1) (10 mL). The reaction mixture was stirred for 2 hours at room temperature. The mixture was concentrated *in vacuo* and the residue was triturated with Et_2O. The resulting white solid was isolated by filtration and washed with Et_2O to yield monomer **(61)** as a white solid (320 mg, 95 %).

Data: R_f = 0.67 (3:1:1 propanol:H_2O:NH_3) UV; υ_{max} (solid, cm^{-1}) 1712 (m), 1690 (s), 1641 (s), 1536 (m), 1480 (w), 1440 (w), 1404 (w), 1357 (w), 12893 (m), 1262 (m), 1194 (s), 1132 (s), 1044 (m), 1018 (w); δ_H (400 MHz, d_6-DMSO) (2:1 Rotamers) 12.92 (3H br s, CO_2**H**), 9.42 (4H, br s, N**H**$_2$), 8.77 (2H, br s, N**H**$_2$), 7.96-8.03 (6H, m, **H**$^{4/5}$), 7.81-7.93 (3H, m, **H**6), 7.72-7.81 (8H, m, **H**$^{1/8}$, N**H**), 7.48-7.57 (6H, m, **H**$^{2/7}$), 7.32-7.39 (1H, m, N**H**), 7.41-7.49 (6H, m, **H**$^{3/6}$), 6.09-6.18 (3H, m, **H**5), 4.95 (4H, s, C**H**$_2$$CO_2$H), 4.76 (2H, s, C**H**$_2$$CO_2$H), 4.50-4.30 (13H, m, NCO_2C**H**$_2$N, C**H**$_2$CH^9, C**H**^{9}), 4.11 (2H, s, NCO_2C**H**$_2$N), 3.60-3.20 (12H, m, NHC**H**$_2$C**H**$_2$N); δ_C (100 MHz, d_6-DMSO) 173.89 (0), 172.63 (0), 170.71 (0), 158.42 (0), 150.77 (0), 149.24 (1), 144.32 (0), 141.19 (0), 128.07 (1), 127.51 (1), 125.57 (1), 120.56 (1), 93.67 (1), 65.98 (2), 65.35 (2), 49.51 (2), 48.20 (2), 47.19 (2), 39.07 (1); LRMS (ES+) *m/z* 492.2 $(M+H)^+$ (100 %); HRMS (ES+) *m/z* 492.1871 $(M+H)^+$; $C_{25}H_{26}N_5O_6$ requires 492.1878.

2-[[(6-Amino-9*H*-9-purinyl)carbonyl](2-[(9*H*-9-fluorenylmethoxy)carbonyl]aminoethyl)amino]acetic acid (63)



Monomer **(62)** (490 mg, 0.68 mmol) was added to a solution of TFA:*m*-cresol (4:1) (10 mL). The reaction mixture was stirred for 2 hours under at room temperature. The mixture was concentrated *in vacuo* and the residue was triturated with Et_2O. The resulting white solid was isolated by filtration and washed with Et_2O to yield monomer **(63)** as a colourless solid (334 mg, 96 %).

Data: R_f = 0.66 (3:1:1 propanol:H_2O:NH_3) UV; υ_{max} (Solid, cm^{-1}) 1710 (s), 1689 (s), 1642 (s), 1536 (m), 1479 (w), 1449 (w), 1405 (w), 1357 (w), 1288 (m), 1260 (m), 1192 (s), 1133 (s), 1104 (m), 1017 (w); δ_H (400 MHz, d_6-DMSO) (2:1 rotamers) 8.61 (6H, br s, N**H**$_2$), 8.44 (2H, br s, **H**8), 8.39 (1H, br s, **H**8), 8.33 (2H, br s, **H**2), 8.32 (1H, br s, **H**2), 7.90-8.03 (6H, m, **H**$^{4/5}$), 7.67-7.78 (6H, m, **H**$^{1/8}$), 7.55-7.63 (2H, m, N**H**), 7.48-7.60 (6H, m, **H**$^{3/6}$), 7.32-7.45 (6H, m, **H**$^{2/7}$), 6.63-6.71 (1H, m, N**H**), 5.41 (4H, s, NC**H**$_2$CO$_2$H), 5.17 (2H, s, NC**H**$_2$CO$_2$H), 4.50-4.30 (13H, m, C**H**$_2$CH^9, C**H**^{9}, NCOC**H**$_2$N), 4.12 (2H, s, NCOC**H**$_2$N), 3.57-3.75 (4H, m, NHCH$_2$C**H**$_2$N), 3.37-3.56 (6H, m, NHC**H**$_2$CH$_2$N, NHCH$_2$C**H**$_2$N), 3.18-3.32 (2H, m, NHC**H**$_2$CH$_2$N); δ_C (100 MHz, d_6-DMSO) 170.43 (0), 169.93 (0), 166.51 (0), 166.04 (0), 156.09 (0), 148.97 (1), 147.53 (0), 143.53 (1), 143.35 (0), 140.41 (0), 127.27 (1), 126.71 (1), 124.76 (1), 119.77 (1) 65.20 (2), 64.56 (2), 48.83 (2), 47.42 (1), 46.41 (2), 43.70 (2); LRMS (ES+) *m/z* 516.1 $(M+H)^+$ (100 %); HRMS (ES+) *m/z* 516.1987 $(M+H)^+$; $C_{26}H_{26}N_3O_5$ requires 516.1990.

7. References

1 R. Franklin and R. G. Gosling, *Nature*, 1953, **172**, 156.
2 J. D. Watson and F. H. C. Crick, *Nature*, 1953, **171**, 737.
3 J. D. Watson, 'Molecular Biology of the Gene', Addison-Wesley, 1976.
4 E. Chargaff, *Fed. Proc.*, 1951, **10**, 654.
5 E. Chargaff, E. Vischer, R. Doniger, C. Green, and F. Misani, *J. Biol. Chem.*, 1951, **177**, 405.
6 A. J. Brookes, *Gene*, 1999, **234**, 177.
7 R. Sachidanandam, D. Weissman, S. C. Schmidt, J. M. Kakol, L. D. Stein, G. Marth, S. Sherry, J. C. Mullikin, B. J. Mortimore, D. L. Willey, et al., *Nature*, 2001, **409**, 928
8 B. W. Kirk, M. Feinsod, R. Favis, R. M. Kliman, and F. Barany, *Nucl. Acids. Res.*, 2002, **30**, 3295.
9 K. Kleppe, E. Ohtsuka, R. Kleppe, R. Molineux, and H. G. Khorana, *J. Mol. Biol.*, 1971, **56**, 341.
10 K. B. Mullis and F. A. Faloona, *Method Enzymol.*, 1987, **155**, 335.
11 K. B. Mullis, *Sci Am*, 1990, 36.
12 R. K. Saiki, D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, and H. A. Erlich, *Science*, 1988, **239**, 487.
13 J. J. Thomson, *Nature*, 1913, **90**, 645.
14 R. J. Cotter, *Anal. Chim. Acta*, 1987, **195**, 45.
15 K. Tanaka, H. Waki, Y. Ido, S. Akita, Y. Yoshida, and T. Yoshida, *Rapid Commun. Mass Spectrom.*, 1988, **2**, 151.
16 M. Karas and F. Hillenkamp, *Anal. Chem.*, 1988, **60**, 2299.
17 M. Karas, D. Bachmann, and F. Hillenkamp, *Anal. Chem.*, 1985, **57**, 2935.
18 M. Gluckmann, A. Pfenninger, R. Kruger, M. Thierolf, M. Karas, V. Horneffer, F. Hillenkamp, and K. Strupat, *Int. J. Mass Spectrom.*, 2001, **210**, 121.
19 R. Zenobi and R. Knochenmuss, *Mass Spectrom. Rev.*, 1998, **17**, 337.
20 R. Knochenmuss and R. Zenobi, *Chem. Rev.*, 2003, **103**, 441.
21 M. Karas and R. Kruger, *Chem. Rev.*, 2003, **103**, 427.
22 Y. F. Zhu, K. L. Lee, T. K., S. L. Allman, N. I. Taranenko, and C. H. Chen, *Rapid Commun. Mass Spectrom.*, 1995, **9**, 1315.
23 W. Stephens, *Phys. Rev.*, 1946, **69**, 691.
24 N. I. Ionov and B. A. Mamyrin, *Zh. Tekh. Fiz.*, 1953, **23**, 2101.
25 J.-C. Tabet and R. J. Cotter, *Int. J. Mass Spectrom. Ion Process.*, 1983, **54**, 151.
26 B. A. Mamyrin, V. I. Karataev, D. V. Shmikk, and V. A. Zagulin, *Sov. Phys. JETP*, 1973, **37**, 45.
27 B. A. Mamyrin, *Int. J. Mass Spectrom. Ion Process.*, 1994, **131**, 1.
28 R. S. Brown and J. J. Lennon, *Anal. Chem.*, 1995, **67**, 1998.
29 S. M. Colby, T. B. King, and J. P. Reilly, *Rapid Commun. Mass Spectrom.*, 1994, **8**, 865.

30 W. C. Willey and I. H. McLaren, *Rev. Sci. Instrum.*, 1955, **26**, 1150.
31 M. L. Vestal, P. Juhasz, and S. A. Martin, *Rapid Commun. Mass Spectrom.*, 1995, **9**, 1044.
32 J. R. Lakowicz, 'Principles of Fluorescence Spectroscopy', Plenum Press, 1983.
33 F. Sanger, *Science*, 1981, **214**, 1205.
34 L. M. Smith, Z. S. Sanders, R. J. Kaiser, P. Hughes, C. Dodd, C. R. Connel, S. Heiner, L. E. Kent, and L. E. Hood, *Nature*, 1986, **321**, 674.
35 J. Wilhelm and A. Pingoud, *Chembiochem*, 2003, **4**, 1120.
36 P. M. Holland, R. D. Abramson, R. Watson, and D. H. Gelfand, *Proc. Natl. Acad. Sci. U. S. A.*, 1991, **88**, 7276.
37 K. J. Livak, S. A. J. Flood, J. Marmaro, W. Giusti, and K. Deetz, *PCR Methods and Applications*, 1995, 357.
38 L. G. Lee, C. R. Connell, and W. Bloch, *Nucleic Acids Res.*, 1993, **21**, 3761.
39 T. Morris, B. Robertson, and M. Gallagher, *J. Clin. Microbiol.*, 1996, **34**, 2933.
40 I. Laurendeau, M. Bahuau, N. Vodovar, C. Larramendy, M. Olivi, I. Bieche, M. Vidaud, and D. Vidaud, *Clin. Chem.*, 1999, **45**, 982.
41 S. Tyagi, U. Landegren, M. Tazi, P. M. Lizardi, and F. R. Kramer, *Proc. Natl. Acad. Sci. USA*, 1996, **93**, 5395.
42 L. G. Kostrikis, S. Tyagi, M. M. Mhlanga, D. D. Ho, and F. R. Kramer, *Science*, 1998, **279**, 1228.
43 I. A. Nazarenko, S. K. Bhatnagar, and R. J. Hohman, *Nucleic Acids Res.*, 1997, **25**, 2516.
44 D. Whitcombe, J. Theaker, S. P. Guy, T. Brown, and S. Little, *Nat. Biotechnol.*, 1999, **17**, 804.
45 A. Solinas, L. J. Brown, C. McKeen, J. M. Mellor, J. T. G. Nicol, N. Thelwell, and T. Brown, *Nucl. Acids. Res.*, 2001, **29**, e96.
46 C. A. Mein, B. J. Barratt, M. G. Dunn, T. Siegmund, A. N. Smith, L. Esposito, S. Nutland, H. E. Stevens, A. J. Wilson, M. S. Phillips, et al., *Genome Res.*, 2000, **10**, 330.
47 Y. Mashima, M. Nagano, T. Funayama, Q. Zhang, T. Egashira, J. Kudho, N. Shimizu, and Y. Oguchi, *Clin. Biochem.*, 2004, **37**, 268.
48 M. Nagano, S. Yamashita, K. Hirano, M. Ito, T. Maruyama, M. Ishihara, Y. Sagehashi, T. Oka, T. Kujiraoka, H. Hattori, et al., *J. Lipid Res.*, 2002, **43**, 1011.
49 K. J. Wu, A. Stedding, and C. H. Becker, *Rapid Commun. Mass Spectrom.*, 1993, **7**, 191.
50 P. A. Limbach, P. F. Crain, and J. A. McCloskey, *Curr. Opin. Biotechnol.*, 1995, **6**, 96.
51 P. A. Limbach, *Mass Spectrom. Rev.*, 1996, **15**, 297.
52 E. Nordoff, F. Kirpekar, and P. Roepstorff, *Mass Spectrom. Rev.*, 1996, **15**, 67.
53 L. Haff, P. Juhasz, S. Martin, M. Roskey, I. Smirnov, W. Stanick, M. Vestal, and K. Waddell, *Analusis*, 1998, **26**, M26.
54 P. F. Crain and J. A. McCloskey, *Curr. Opin. Biotechnol.*, 1998, **9**, 25.
55 I. G. Gut, *Hum. Mutat.*, 2004, **23**, 437.

56 E. Nordhoff, A. Ingendoh, P. Cramer, A. Overberg, B. Stahl, M. Karas, and F. Hillenkamp, *Rapid Commun. Mass Spectrom.*, 1992, **6**.

57 G. J. Langley, J. M. Herniman, N. L. Davies, and T. Brown, *Rapid Commun. Mass Spectrom.*, 1999, **13**, 1717.

58 U. Pieles, W. Zurcher, M. Schar, and H. Moser, *Nucl. Acids. Res.*, 1993, **21**, 3191.

59 C. Liu, Q. Wu, A. C. Harms, and R. D. Smith, *Anal. Chem.*, 1996, **68**, 3295

60 C. L. Liu, D. C. Muddiman, K. Q. Tang, and R. D. Smith, *J. Mass Spectrom.*, 1997, **32**, 425.

61 J. C. Hannis and D. C. Muddiman, *Rapid Commun. Mass Spectrom.*, 1999, **13**, 954.

62 D. S. Wunschel, K. F. Fox, A. Fox, J. E. Bruce, D. C. Muddiman, and R. D. Smith, *Rapid Commun. Mass Spectrom.*, 1996, **10**, 29.

63 D. S. Wunschel, D. C. Muddiman, K. F. Fox, A. Fox, and R. D. Smith, *Anal. Chem.*, 1998, **70**, 1203.

64 J. T. Stults and J. C. Marsters, *Rapid Commun. Mass Spectrom.*, 1991, **5**, 359.

65 D. P. Little, A. Braun, B. DarnhoferDemar, A. Frilling, Y. Z. Li, R. T. McIver, and H. Koster, *J. Mol. Med.*, 1997, **75**, 745.

66 P. L. Ross and P. Belgrader, *Anal. Chem.*, 1997, **69**, 3966.

67 E. Nordhoff, C. Luebbert, G. Thiele, V. Heiser, and H. Lehrach, *Nucl. Acids. Res.*, 2000, **28**, e86.

68 L. A. Haff and I. P. Smirnov, *Genome Res.*, 1997, **7**, 378.

69 J. A. Ragas, T. A. Simmons, and P. A. Limbach, *Analyst*, 2000, **125**, 575.

70 N. I. Taranenko, V. V. Golovlev, S. L. Allman, N. V. Taranenko, C. H. Chen, J. Hong, and L. Y. Chang, *Rapid Commun. Mass Spectrom.*, 1998, **12**, 413.

71 M. J. Doktycz, Hurst G. B., Habibigoudarzi S., Mcluckey S. A., Tang K., Chen C. H., Uziel M., Jacobson K. B., Woychik R. P., and Buchanan M. V., *Anal. Biochem.*, 1995, **230**, 205.

72 P. L. Ross, P. A. Davis, and P. Belgrader, *Anal. Chem.*, 1998, **70**, 2067

73 F. Sanger, S. Nicklen, and A. R. Coulson, *PNAS*, 1977, **74**.

74 N. I. Taranenko, S. L. Allman, V. V. Golovlev, N. V. Taranenko, N. R. Isola, and C. H. Chen, *Nucl. Acids. Res.*, 1998, **26**, 2488.

75 J. R. Edwards, Y. Itagaki, and J. Y. Ju, *Nucl. Acids. Res.*, 2001, **29**, e104.

76 F. Kirpekar, E. Nordhoff, L. K. Larsen, K. Kristiansen, P. Roepstorff, and F. Hillenkamp, *Nucl. Acids. Res.*, 1998, **26**, 2554.

77 I. P. Smirnov, M. T. Roskey, P. Juhasz, E. J. Takach, S. A. Martin, and L. A. Haff, *Anal. Biochem.*, 1996, **238**, 19.

78 C. M. Bentzley, M. V. Johnston, and B. S. Larsen, *Anal. Biochem.*, 1998, **258**, 31.

79 L.-K. Zhang and M. L. Gross, *J. Am. Soc. Mass Spectrom.*, 2000, **11**, 854.

80 Y.-H. Liu, J. Bai, X. Liang, D. M. Lubman, and P. J. Venta, *Anal. Chem.*, 1995, **67**, 3482.

81 S. Sauer and I. G. Gut, *J. Chromatogr., B*, 2002, **782**, 73.
82 J. Tost and I. G. Gut, *Mass Spectrom. Rev.*, 2002, **21**, 388.
83 K. Tang, D. Opalsky, K. Abel, D. van den Boom, P. Yip, G. Del Mistro, A. Braun, and C. R. Cantor, *Int. J. Mass Spectrom.*, 2003, **226**, 37.
84 Z. J. Meng, T. A. Simmons-Willis, and P. A. Limbach, *Biomol. Eng.*, 2004, **21**, 1.
85 A. Braun, D. P. Little, and H. Koster, *Clin. Chem.*, 1997, **43**, 1151.
86 A. Braun, D. P. Little, D. Reuter, B. Muller-Mysok, and H. Koster, *Genomics*, 1997, **46**, 18.
87 D. P. Little, T. J. Cornish, M. J. Odonnell, A. Braun, R. J. Cotter, and H. Koster, *Anal. Chem.*, 1997, **69**, 4540.
88 K. Tang, P. Oeth, S. Kammerer, M. F. Denissenko, J. Ekblom, C. Jurinke, D. van den Boom, A. Braun, and C. R. Cantor, *J. Proteome Res.*, 2004, **3**, 218.
89 K. Tang, D. J. Fu, D. Julien, A. Braun, C. R. Cantor, and H. Koster, *PNAS*, 1999, **96**, 10016.
90 J. Li, J. M. Butler, Y. P. Tan, H. Lin, S. Royer, L. Ohler, T. A. Shaler, J. M. Hunter, D. J. Pollart, J. A. Monforte, et al., *Electrophoresis*, 1999, **20**, 1258.
91 P. Ross, L. Hall, I. Smirnov, and L. Haff, *Nat. Biotechnol.*, 1998, **16**, 1347.
92 A. Harksen, P. M. Ueland, H. Refsum, and K. Meyer, *Clin. Chem.*, 1999, **45**, 1157.
93 K. Meyer, A. Fredriksen, and P. M. Ueland, *Clin. Chem.*, 2004, **50**, 391.
94 H. Y. Yang, H. J. Wang, H. Wang, Y. Cai, G. Q. Zhou, F. C. He, and X. H. Qian, *Anal. Biochem.*, 2003, **314**, 54.
95 L. A. Haff and I. P. Smirnov, *Nucl. Acids. Res.*, 1997, **25**, 3749.
96 Z. D. Fei, T. Ono, and L. M. Smith, *Nucl. Acids. Res.*, 1998, **26**, 2827.
97 S. Kim, S. Shi, T. Bonome, M. E. Ulz, J. R. Edwards, H. Fodstad, J. J. Russo, and J. Ju, *Anal. Biochem.*, 2003, **316**, 251.
98 F. Abdi, E. M. Bradbury, N. Doggett, and X. Chen, *Nucl. Acids. Res.*, 2001, **29**, e61.
99 K. Hung, X. Sun, H. Ding, M. Kalafatis, P. Simioni, and B. Guo, *Blood Coagul Fibrinolysis.*, 2002, **13**, 117.
100 X. Sun, H. Ding, K. Hung, and B. Guo, *Nucl. Acids. Res.*, 2000, **28**, e68.
101 S. Sauer, D. Lechner, K. Berlin, C. Plancon, A. Heuermann, H. Lehrach, and I. G. Gut, *Nucl. Acids. Res.*, 2000, **28**, e100.
102 S. Sauer, D. Lechner, K. Berlin, H. Lehrach, J. L. Escary, N. Fox, and I. G. Gut, *Nucl. Acids. Res.*, 2000, **28**, e13.
103 I. G. Gut, W. A. Jeffery, D. J. C. Pappin, and S. Beck, *Rapid Commun. Mass Spectrom.*, 1997, **11**, 43/50.
104 I. Gut and S. Beck, *Nucl. Acids. Res.*, 1995, **23**, 1367.
105 S. Sauer, H. Lehrach, and R. Reinhardt, *Nucl. Acids. Res.*, 2003, **31**, e63.
106 S. Sauer and I. G. Gut, *Rapid Commun. Mass Spectrom.*, 2003, **17**, 1265.

107 V. Lyamichev, A. L. Mast, J. G. Hall, J. R. Prudent, M. W. Kaiser, T. Takova, R. W. Kwiatkowski, T. J. Sander, M. de Arruda, D. A. Arco, et al., *Nat. Biotechnol.*, 1999, **17**, 292.

108 J. G. Hall, P. S. Eis, S. M. Law, L. P. Reynaldo, J. R. Prudent, D. J. Marshall, H. T. Allawi, A. L. Mast, J. E. Dahlberg, R. W. Kwiatkowski, et al., *PNAS*, 2000, **97**, 8272.

109 M. Olivier, L. M. Chuang, M. S. Chang, Y. T. Chen, D. Pei, K. Ranade, A. de Witte, J. Allen, N. Tran, D. Curb, et al., *Nucl. Acids. Res.*, 2002, **30**, e53.

110 T. J. Griffin and L. M. Smith, *Anal. Chem.*, 2000, **72**, 3298.

111 T. J. Griffin, J. G. Hall, J. R. Prudent, and L. M. Smith, *PNAS*, 1999, **96**, 6301.

112 S. Kwok and R. Higuchi, *Nature*, 1989, **339**, 237.

113 P. L. Ross, K. Lee, and P. Belgrader, *Anal. Chem.*, 1997, **69**, 4197.

114 P. JiangBaucom, J. E. Girard, J. Butler, and P. Belgrader, *Anal. Chem.*, 1997, **69**, 4894.

115 T. J. Griffin, W. Tang, and L. M. Smith, *Nat. Biotechnol.*, 1997, **15**, 1368.

116 M. Egholm, O. Buchardt, P. E. Nielsen, and R. H. Berg, *J. Amer. Chem. Soc.*, 1992, **114**, 1895.

117 P. E. Nielsen, M. Egholm, R. H. Berg, and O. Buchardt, *Science*, 1991, **254**, 1497.

118 M. Egholm, P. E. Nielsen, O. Buchardt, and R. H. Berg, *J. Am. Chem. Soc.*, 1992, **114**, 9677.

119 S. K. Kim, P. E. Nielsen, M. Egholm, O. Buchardt, R. H. Berg, and B. Norden, *J. Am. Chem. Soc.*, 1993, **115**, 6477.

120 S. Tomac, M. Sarkar, T. Ratilainen, P. Wittung, P. E. Nielsen, B. Norden, and A. Graslund, *J. Am. Chem. Soc.*, 1996, **118**, 5544.

121 F. P. Schwarz, S. Robinson, and J. M. Butler, *Nucl. Acids. Res.*, 1999, **27**, 4792.

122 M. Egholm, O. Buchardt, L. Christensen, C. Behrens, S. M. Freier, D. A. Driver, R. H. Berg, S. K. Kim, B. Norden, and P. E. Nielsen, *Nature*, 1993, **365**, 556.

123 P. E. Nielsen, *Curr. Opin. Struct. Biol.*, 1999, **9**, 353.

124 P. E. Nielsen, *Curr. Opin. Biotechnol.*, 1999, **10**, 71.

125 H. Stender, M. Fiandaca, J. J. Hyldig-Nielsen, and J. Coull, *J. Microbiological Methods*, 2001.

126 B. Hyrup and P. E. Nielsen, *Bioorg. Med. Chem.*, 1996, **4**, 5.

127 M. C. Griffith, L. M. Risen, M. J. Greig, E. A. Lesnik, K. G. Sprankle, R. H. Griffey, J. S. Kiely, and S. M. Freier, *J. Amer. Chem. Soc.*, 1995, **117**, 831

128 S. Sforza, T. Tedeschi, R. Corradini, A. Dossena, and R. Marchelli, *Chem. Commun.*, 2003, 1102.

129 J. W. Flora and D. C. Muddiman, *J. Am. Soc. Mass Spectrom.*, 2001, **12**, 805.

130 J. W. Flora, D. D. Shillady, and D. C. Muddiman, *J. Am. Soc. Mass Spectrom.*, 2000, **11**, 615.

131 J. W. Flora and D. C. Muddiman, *Rapid Commun. Mass Spectrom.*, 1998, **12**, 759.
132 J. M. Butler, P. JiangBaucom, M. Huang, P. Belgrader, and J. Girard, *Anal. Chem.*, 1996, **68**, 3283.
133 J. M. Rommens, M. C. Iannuzzi, B.-s. Kerem, M. L. Drumm, G. Melmer, M. Dean, R. Rozmahel, J. L. Cole, D. Kennedy, N. Hidaka, et al., *Science*, 1989, **245**, 1059.
134 F. S. Collins, *Science*, 1992, **256**, 774.
135 'PNA Chemistry for the Expedite 8900 Nucleic acid synthesis system', PerSeptive Biosystems, 1998.
136 G. L. Igloi, *PNAS*, 1998, **95**, 8562.
137 T. A. Early, J. Olmsted, D. R. Kearns, and A. G. Lezius, *Nucl. Acids. Res.*, 1978, **5**, 1955–1970.
138 B. D. Gildea, S. Casey, J. MacNeill, H. Perry-O'Keefe, D. Sorensen, and J. M. Coull, *Tetrahedron Lett.*, 1998, **39**, 7255.
139 S. Hahner, J. Olejnik, H. C. Ludemann, E. Krzymanska-Olejnik, F. Hillenkamp, and K. J. Rothschild, *Biomol. Eng.*, 1999, **16**, 127.
140 X. P. Bai, Z. M. Li, S. Kim, H. D. Ruparel, N. J. Turro, and J. Y. Ju, *Biochemistry*, 2003, **42**, 18.
141 X. P. Bai, S. B. Kim, Z. M. Li, N. J. Turro, and J. Y. Ju, *Nucl. Acids. Res.*, 2004, **32**, 535.
142 C. P. Holmes and D. G. Jones, *J. Org. Chem.*, 1995, **60**, 2318.
143 K. Burgess, C. I. Martinez, D. H. Russell, H. Shin, and A. J. Zhang, *J. Org. Chem.*, 1997, **62**, 5662.
144 C. P. Holmes, *J. Org. Chem.*, 1997, **62**, 2370.
145 M. Bartlet-Jones, W. A. Jeffery, H. H. F., and D. J. C. Pappin, *Rapid Commun. Mass Spectrom.*, 1994, **8**, 737.
146 B. Spengler, F. Luetzenkirchen, S. Metzger, P. Chaurand, R. Kaufmann, W. Jeffery, M. Bartlet-Jones, and D. J. C. Pappin, *Int. J. Mass Spectrom. Ion Process.*, 1997, **169-170**, 127.
147 K. D. Roth, Z.-H. Huang, N. Sadagopan, and J. T. Watson, *Mass Spectrom. Rev.*, 1998, **17**, 255.
148 N. A. Stewart, V. T. Pham, C. T. Choma, and H. Kaplan, *Rapid Commun. Mass Spectrom.*, 2002, **16**, 1448.
149 N. Thelwell, S. Millington, A. Solinas, J. Booth, and T. Brown, *Nucl. Acids. Res.*, 2000, **28**, 3752.
150 H. E. Moser and P. B. Dervan, *Science*, 1987, **238**, 645.
151 N. T. Thuong and C. Hélène, *Angew, Chemie. Int. Ed. Engl.*, 1993, **32**, 666.
152 K. R. Fox, *Curr. Med. Chem.*, 2000, 17.
153 D. Guianvarc'h, R. Benhida, J.-L. Fourrey, R. Maurisse, and J.-S. Sun, *Chem Commun*, 2001, 1814.
154 P. E. Nielsen and M. Egholm, *Bioorg. Med. Chem.*, 2001, **9**, 2429.
155 L. Betts, J. A. Josey, J. M. Veal, and S. R. Jordan, *Science*, 1995, **270**, 1838.
156 H. Challa and S. A. Woski, *Tetrahedron Lett.*, 1999, **40**, 419.
157 H. Challa and S. A. Woski, *Tetrahedron Lett.*, 1999, **40**, 8333.

158 B. Armitage, T. Koch, H. Frydenlund, H. Orum, and G. Schuster, *Nucl. Acids. Res.*, 1998, **26**, 715.
159 S. Sforza, R. Corradini, G. Galaverna, A. Dossena, and R. Marchelli, *Minerva Biotechnol.*, 1999, **11**, 163.
160 S. A. Thompson, J. A. Hosey, R. Cadilla, M. D. Gaul, C. F. Hassman, M. J. Luzzio, A. J. Pipe, K. L. Reed, D. J. Ricca, R. W. Wiethe, et al., *Tetrahedron*, 1995, **51**, 6179.
161 G. Breipohl, D. W. Will, A. Peyman, and E. Uhlmann, *Tetrahedron*, 1997, **53**, 14671.
162 E. P. Heimer, H. E. Gallo-torres, A. M. Felix, M. Ahmad, T. J. Lambros, F. Scheidl, and J. Meienhofer, *Int. J. Pept. Protein Res.*, 1984, **23**, 203.
163 Salvi, Jean-Paul, N. Walchshofer, and J. Paris, *Tetrahedron Lett.*, 1994, **35**, 1181.
164 D. W. Will, G. Breiphol, D. Langner, J. Knolle, and E. Uhlmann, *Tetrahedron*, 1995, **51**, 12069.
165 G. Gellerman, A. Elgavi, Y. Salitra, and I. Kramer, *J. Pept. Res.*, 2001, **57**, 277.
166 S. V. Ley, J. Norman, W. P. Grifith, and S. P. Marsden, *Synthesis*, 1994, 639.
167 T. Fukuyama, C. K. Jow, and M. Cheung, *Tetrahedron Lett.*, 1995, **36**, 6373.
168 T. Fukuyama, M. Cheung, C. K. Jow, Y. Hidai, and T. Kan, *Tetrahedron Lett.*, 1997, **38**, 5831.
169 S. J. Barry, R. M. Carr, S. J. Lane, W. J. Leavens, S. Monte, and I. Waterhouse, *Rapid Commun. Mass Spectrom.*, 2003, **17**, 603.
170 Bernatowicz, Michael S., S. B. Daniels, and H. Koster, *Tetrahedron Lett.*, 1989, **30**, 4645.
171 G. Kovacs, Z. Timar, Z. Kupihar, Z. Kele, and L. Kovacs, *J. Chem. Soc.-Perkin Trans. 1*, 2002, 1266.
172 Z. Timar, L. Kovacs, G. Kovacs, and Z. Schmel, *J. Chem. Soc.-Perkin Trans. 1*, 2000, 19.
173 M. Rodriguez, M. Llinares, S. Doulut, A. Heitz, and J. Martinez, *Tetrahedron Lett.*, 1991, **32**, 923.