# TOWARDS THE SYNTHESIS OF FERROCENE NUCLEIC ACIDS

**By Claire Elizabeth Lydon**

A Thesis Submitted to The University of Birmingham

for the Degree of MRes Materials Chemistry and Nanochemistry

Department of Chemistry
The University of Birmingham
September 2008

# ABSTRACT

The fields of DNA chemistry and electrochemical sensing were combined in an attempt to synthesise a ferrocene modified nucleobase, which could replace two nucleotides in a DNA sequence.

Two modified ferrocene monomers were attempted to be incorporated into oligonucleotide strands, using solid state phosphoramidite DNA synthesis techniques, and characterised using MALDI-ToF Mass Spectrometry. CD and $T_m$ studies were also carried out.

Synthesis of a modified ferrocene, with two functional groups on each of the cyclopentadienyl rings (four in total) was attempted, using a protected alcohol, to allow the necessary functionalisation for a ferrocene nucleobase.

Progress was made towards synthesising a precursor of the ferrocene nucleobase, using a procedure for the preparation of a chiral bis-amine.

# ACKNOWLEGEMENTS

Thanks to my supervisor, Dr. James Tucker for his guidance in the project and to the members of his research group who have assisted me in my research and lab work over the past year. I would also like to thank Dr. Joseph Vyle and his research group in the department of chemistry at Queens University Belfast, for their help in attempting to synthesise the modified oligonucleotide strands. Dr. Neil Spencer, Peter Ashton and Nick May of NMR and Mass Spectrometry facilities providing the majority of my characterisation data have been particularly useful.

Special thanks also go to Jean-Louis Duprey, who has been of great help in both practical aspects of the project and in the writing up process.

# TABLE OF CONTENTS

# TABLE OF FIGURES

# TABLE OF TABLES

# ABBREVIATIONS

| | |
|---|---|
| A | Adenine |
| BTT | 5-Benzylthio-1-H-tetrazole |
| C | Cytosine |
| CD | Circular Dichroism |
| $CD_3CN$ | Deuterated Acetonitrile |
| $CDCl_3$ | Deuterated chloroform |
| Cp | Cyclopentadiene ring |
| CPG | Controlled Pore Glass |
| d | Doublet |
| DCM | Dichloromethane |
| dd | Doublet of doublets |
| DIPEA | N,N-Diisopropylethylamine |
| DMAP | Dimethylaminopyridine |
| DMF | Dimethylformamide |
| DMT | Dimethoxytrityil |
| DNA | Deoxyribonucleic acid |
| ES | Electrospray |
| Fc | Ferrocene |
| G | Guanine |
| HPLC | High Performance Liquid Chromatography |
| M | Molarity |

| | |
|---|---|
| m | Multiplet |
| MALDI-ToF | Matrix Assisted Laser Desorption Ionisation- Time of Flight |
| MHz | Megahertz |
| mmol | Millimole |
| nm | Nanometre |
| NMR | Nuclear magnetic Resonance |
| Ppm | Parts per million |
| q | Quartet |
| R.T. | Room Temperature |
| RNA | Ribonucleic Acid |
| s | Singlet |
| t | Triplet |
| T | Thymine |
| TBDMS | Tertbutyldimethylsilyl |
| TEA | Triethylamine |
| THF | Tetrahydrofuran |
| TLC | Thin layer chromatography |
| $T_m$ | Melting temperature |
| U | Uridine |
| UV/Vis | Ultra-violet/ visible |
| $\delta$ | Chemical Shift |
| μl | microlitre |
| μM | micromolar |

# 1 INTRODUCTION

Ferrocene can be (simply) defined as:

An "iron sandwich," in which the iron is the filling and the slices of bread are the aromatic cyclopentadienyl rings

\- Geoffrey Wilkinson

## *1.1 Background*

Ferrocene was first discovered in 1951 (called dicyclopentadienyl iron) and reported separately by the groups of Kealy and Pauson (who published the first paper on it)[1] and Miller, Treboth and Tremaine.[2] It was the first of the metallocenes to be synthesised. The structure was identified as a so called "sandwich" compound in 1952 by Fischer (from preliminary X-Ray data)[3] and by Wilkinson (using IR data and diamagnetism),[4] who proposed calling the compound ferrocene. It was hypothesised that the cyclopentadienyl (Cp) rings of the ferrocene had aromatic character and sat above and below the iron atom. This proposed structure was confirmed by X-Ray Crystallography by Orgel and Dunitz,[5] Eiland and Repinsky[6] and Moffitt.[7] It also became apparent that there was indeed aromaticity in the compound when aromatic electrophilic substitution reactions, such as Friedel-Crafts acylation, were successfully carried out.[8] In the solid state, the Cp rings were found to be most stable in the anti-prismatic or staggered form, unlike the other metallocenes of the same group (osmocene and ruthocene), which sit in the prismatic, or eclipsed conformation (**Figure 1**). It has been suggested that this is because there is a smaller distance between the cp rings in ferrocene and therefore more interannular repulsion. The staggered form allows the non-bonded interactions between heteroannular carbons to be minimised.[9] However, the energy barrier for rotation of the rings is very low and so they may rotate freely around the five-fold symmetry axis (allowing for both eclipsed and staggered forms to be observed).

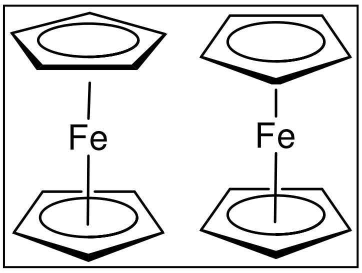

**Figure 1**. Ferrocene in the staggered (left) and eclipsed (right) forms

The Cp rings are parallel plane regular pentagons and held together by $sp^2$ σ- bonds, which are slightly bent. The C 2p orbitals can form molecular orbitals extending around the ring (delocalisation), which can then be involved in π- bonding to the metal.

Ferrocene is remarkably stable for an organometallic substance: with heating up to 470 °C and to acids and alkalis,[10] which was unprecedented at the time of discovery (it had been typically thought that transition metals would not form stable bonds to hydrocarbons). This stability is attributed to the aromaticity of the Cp rings. There are six electrons in the π-orbitals of each ring, thus giving aromatic character (as in benzene). The iron atom in ferrocene is considered to be in the +2 oxidation state, and the Cp rings assumed to carry a single negative charge each, giving them six electrons per ring. This brings the total electron count to 18 (krypton noble gas configuration), which contributes to the stability and accounts for the diamagnetic behaviour. The Cp rings can be said to occupy three co-ordination sites of the iron atom, and this allows for octahedral geometry.[11]

The iron atom is bonded to the Cp *via* its d-orbitals and the π-orbitals on the rings. The C-H bonds appear as a singlet stretch in the IR spectrum, which indicates that they are all equivalent, as in benzene. This means that there are no carbon-carbon double bonds, as there would be in cyclopentadienyl, which can also be seen in the C-C bond lengths, which are all 1.40 Å.[11]

It can be seen that ferrocene has some interesting properties; indeed, it is still a much-studied molecule and earned the people who first elucidated the structure a Nobel Prize.[12] In particular, it has some interesting and useful redox chemistry; the iron atom can be oxidised from the +2 to +3 state in a reversible one electron transfer process. The blue cation is known as a ferricenium ion.[13]

## *1.2 Ferrocene in Biology and Medicine*

As it is so stable in such varied conditions, ferrocene can be used for such varied purposes as electrochemical sensing, polymers and materials, in the field of medicine and as catalysts for organic reactions. Its stability in aqueous, aerobic media and the large number of derivatives that are easily accessible has prompted speculation about the potential for biological applications, and the field of bioorganometallic chemistry has flourished in recent years.[14]

It is thought that derivatives of ferrocene could be used in the treatment of various diseases, including malaria (ferroquine, currently in clinical trials)[15] and cancer (anti-tumour agents,[16] such as an analogue of tamoxifen).[17]

Another area in which ferrocene could be utilised is in biomolecule sensing. It is desirable to undertake biomolecule sensing in order to make detection and understanding of problems in the body more simple and effective, which would allow them to be treated earlier and more successfully. An example of this is the detection of glucose to monitor blood sugar levels in diabetics with glucose sensors.[12] The more accurate the sensor, the earlier a problem can be detected, and the greater the control the patient has over their condition. Attaching ferrocene to proteins, or DNA would allow defects in their make up to be identified and the molecule could be tracked by electrochemical sensing to study binding. This could lead to a better understanding of genetic defects, as well as practical uses in monitoring bodily processes, as described above.

## *1.3 The Structure and Properties of DNA*

DNA is a polyanionic double helix, consisting of two polymeric strands (**Figure 2**).

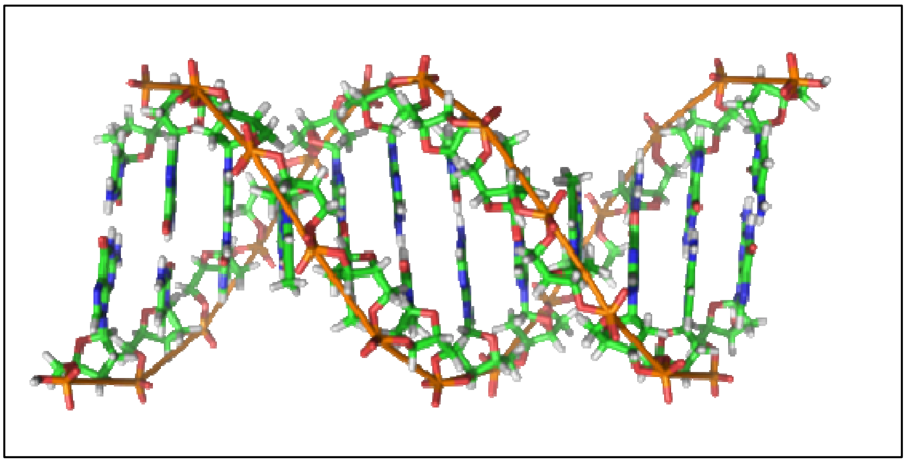

**Figure 2**. B DNA – the most common form encountered.

**Taken From http://commons.wikimedia.org/wiki/Image:A-DNA,_B-DNA_and_Z-DNA.png**

Each strand consists of repeating units called nucleotides and runs in either the 3’ or 5’ (numbering of carbons of the sugar) direction. The nucleotides are composed of a 2-deoxyribose sugar, covalently linked to a negatively charged phosphate group, which connects the sugars, acting as a backbone, and a pyrimidine or purine derived base, also attached to the sugar. There are four bases available for attachment to the sugar and they form complementary hydrogen bonds to bases on another strand to give the familiar two stranded duplex (**Figure 3**). These bases are Adenine, Cytosine, Thymine and Guanine (A, C, T and G).

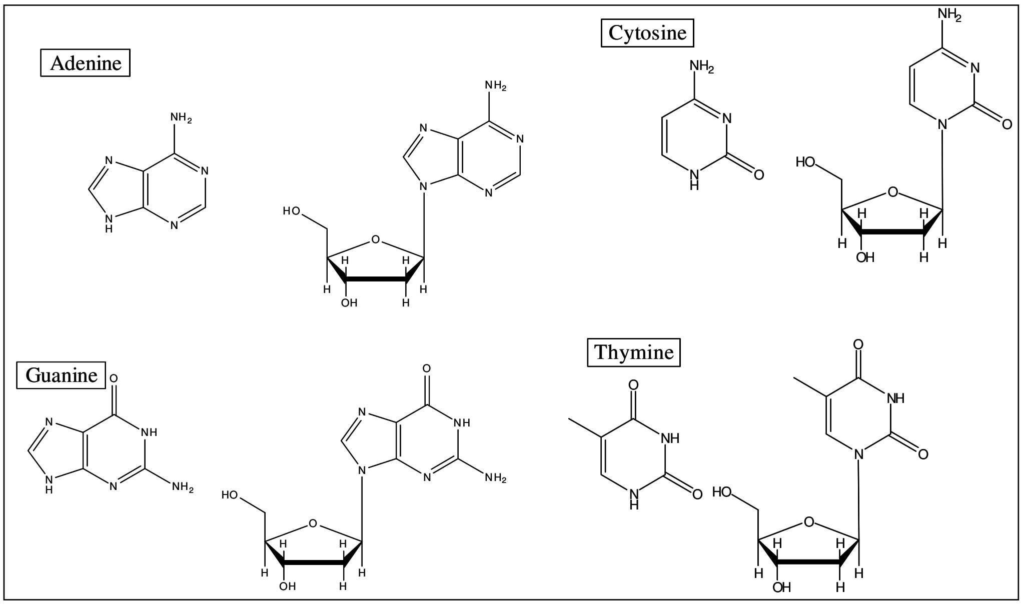

**Figure 3.** The purine bases A and G and the respective nucleotides (left) and pyrimidines C and T with respective nucleotides.

A is complementary to T, forming two hydrogen bonds, and C to G, forming three hydrogen bonds (**Figure 4**). G and A are purines with two heterocyclic rings (one five-membered and one six-membered) and C and T are pyrimidines, with a six membered heterocyclic ring structure.

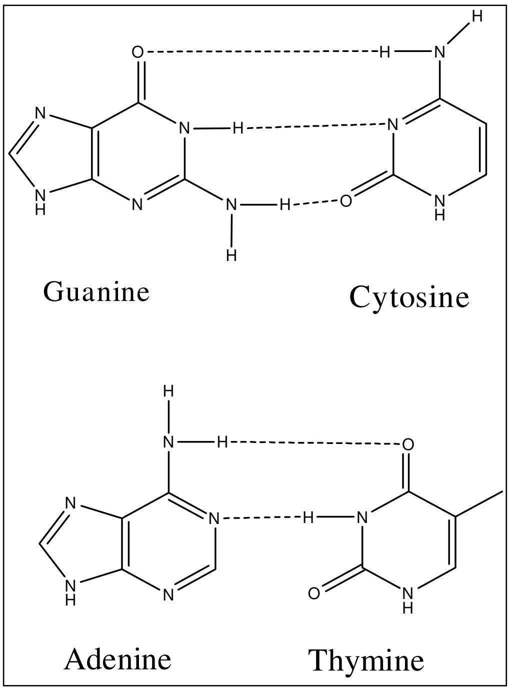

**Figure 4**. Non-covalent bonding in the DNA base pairings CG and AT

As well as interacting across the strands, the bases also engage in base stacking behaviour, with the rings all lying with the flat faces parallel to each other. As the bases consist of aromatic rings, non-covalent π-π interactions can occur, adding to the stability of the duplex.

The spacing between the stacking base pairs in B-DNA (the right handed helix most commonly encountered) is 3.4 Å, with ten base pairs per helical turn.[18]

## *1.4 DNA Sensing*

There has been much research aimed at sensing defects in DNA in recent years. This is because it is thought that the ability to detect the likelihood of inheriting certain diseases could lead to a more personalised healthcare plan, which in turn could increase the probability of a successful outcome to any treatment. This has been very much in the public consciousness recently, with geneticists discovering genes that contribute towards the susceptibility of an individual to certain health problems, such as heart disease, or of their tolerance to certain therapeutic drugs.

Molecular sensors rely on a specific recognition event to detect the desired target and give a measurable signal. Thus, the most important things to include in design of biosensors are a molecular recognition layer and a signal transducer, which can be coupled to an appropriate readout device.[19]

The previously mentioned redox properties of ferrocene (section 1.1) have been utilised for electrochemical sensing purposes with respect to DNA; conjugating ferrocene with oligonucleotides to detect genetic mutations, as elaborated below, is a much researched area. The mutations that could be investigated include deletions (removal of a base), insertions (extra bases on the DNA strand) or polymorphisms (replacement of one base with another), the presence of which can result in diseases which can markedly affect quality of life.

Single Nucleotide Polymorphisms (SNPs) are the most common type of mutation, and identification of these may help to locate particular disease risk factors, such as heart disease or diabetes in individuals. This in turn could lead to personalised medical care, dependent on genetics. SNPs can also be a direct cause of genetic disorders. Examples include cystic fibrosis (CF), where the mutation (of which many, including polymorphisms, can cause the disease) causes misfolding of protein channels which carry chloride ions across cell membranes[20] and thalassemia, which alters the red blood cells in the body, making them less able to take up iron.[21] If the mutations can be detected, then in theory a method of treatment can be evolved.

There are various methods, either utilising enzymes or not, that can be used to determine whether any genetic mutations are present.[22] The design of techniques to detect mutations can be approached from the biology or chemistry standpoint. Much of the research being done from a chemical, as opposed to biological, perspective utilises fluorescent and electrochemical methods. Many of these are based on the fact that bonding of the complementary base pairs between the DNA strands is very stable. The complementary base pairs hold the two strands of DNA together, to make the familiar duplex structure (As seen in **Figure 2** and described above). If the complementary base is replaced by a non-complementary one, the bonding in the duplex is different and a change in signal from whatever type of sensor is being used can be observed, although some of the mismatches are only a little less stable than the complementary pair. The stability of the DNA is measured by studying the melting temperature ($T_m$) of the duplex. The less stable the DNA (with fewer non-covalent interactions), the lower the $T_m$.

There are numerous ways of detecting mismatches electrochemically. One example involves disruption of base stacking in DNA, which in turn reduces the efficiency of charge transport through the DNA, giving a reduced signal.[23] There are five main types of electrochemical sensing platforms. These are direct DNA electrochemistry, indirect DNA electrochemistry, DNA-specific redox indicator detection, nanoparticle-based electrochemistry amplification and DNA-mediated charge transport.[19]

The advantages of using an electrochemical sensor are that they are inexpensive (due to being conducted on a variety of electrode substrates), easy to use and very sensitive (to femtomole concentrations). This is especially true if a fully digital "on-off" signal response can be engineered. The fluorescence based optical devices are also very sensitive and allow parallel analysis of target DNA sequences, but are rather expensive techniques, which require numerical algorithms to interpret the data, possibly limiting the methods to use in research.[14]

The end aim is to have a "lab-on-a-chip" device, with an array of electrodes. This has already been established with fluorescence techniques, and is being researched as regards electrochemical methods. There is currently a 14 microelectrode disposable chip being developed, based on thiolated capture probes and gold microelectrodes.[14]

The idea behind having a lab on a chip is to have many ferrocene probes clustered together onto a single surface to create something akin to a microchip, which can be used in diagnostics for quick, simple and effective DNA testing. For this to be as simple as possible, the detection methods need to be very sensitive, so that amplification of the DNA (by PCR) is not required beforehand. The results must also be reproducible, and ideally analysed in biological media, as this is where DNA in its native state is found.

## *1.5 Ferrocene and DNA Sensing*

Conjugation of ferrocene derivatives to DNA is a relatively immature field, compared to that of protein and amino acid conjugation with ferrocene, but it is one that has the potential to produce accurate and simple gene sensors. This could allow mutations on DNA to be detected, which would indicate the presence of inherited diseases.[14]

Of the five methods mentioned above for electrochemically detecting mismatches, two involve using the electrochemical properties of DNA, one tagging the DNA with a redox active molecule, one uses nanoparticles and the last an intercalator. The intercalator attaches to the duplex *via* the major or minor groove, but the ferrocene moiety is not used to tag the DNA sequence, like the redox indicator type sensor.

A direct DNA sensor utilises the electrochemical properties of guanine (it can be electrochemically oxidised).[24] However, these require high potentials, and as such, give a high background current.

Indirect sensors use electrochemical mediators, such as polypyridyl ruthenium and osmium complexes. In this method, the reduced metal complexes are oxidised and then come into contact with the purine DNA bases, which can reduce the complex. The signal shows how much guanine is available for oxidation.[19]

The ferrocene can be incorporated by attachment of a functionalised molecule to a nucleotide, which is then integrated into an oligonucleotide strand. This is also known as a hybridisation assay (**Figure 5**). The modified nucleotide generally attaches at the end of the DNA strand (probe or target) by solid state reactions or Polymerase Chain Reactions (PCR).[25] This is a DNA-specific redox indicator type sensor, an early example of which is seen in the work of Ihara and colleagues from 1996.[26] This particular sensor had a sensitivity of femtomolar concentrations to DNA and RNA strands.

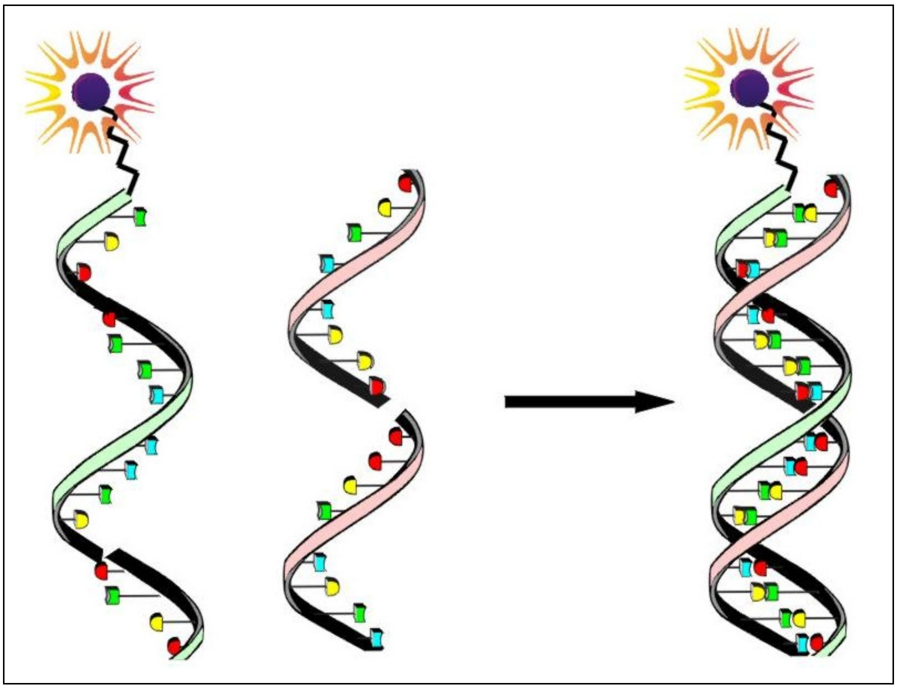

**Figure 5**. Example of a hybridisation assay of a 15-mer oligonucleotide.

Taken From**:** **http://www.iom-berlin.de/html/eng/produkte/reader/2.04.1.4-nanoscan.htm**

Another variation of this detection method is the “sandwich assay,” using three DNA strands. Here, a capture probe is immobilised onto a gold surface, and the target DNA will only bind if the complementary DNA sequence is present. The target DNA possesses a sequence complementary to that of the ferrocene probe. This gives a ferrocene labelled target DNA. If the target binds to the immobilised probe, the ferrocene is brought in close proximity to the electrode and is oxidised, giving a significant peak, as shown by **Figure 6**. A mismatch results in weaker binding, and a smaller peak is observed. This is because there are fewer ferrocene molecules being held close to the electrode.[14] The redox label is attached to a synthetic sequence designed to bind an overhanging section of the probe-target sequence.[19]

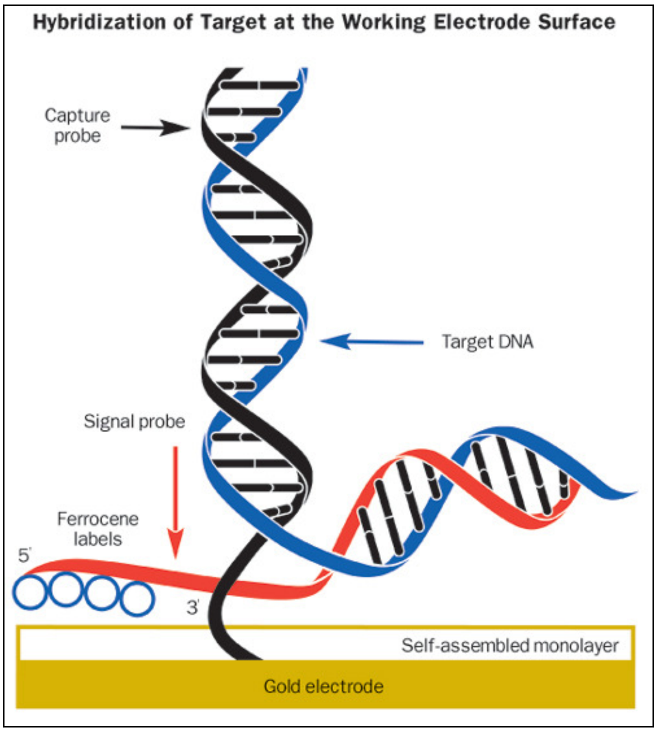

**Figure 6.** Schematic of a sandwich assay, as used in an IVD technology eSensor microarray.

Taken From**:** **www.devicelink.com/ivdt/archive/08/05/007.html**

Nanoparticle amplification involves depositing silver metal onto gold nanoparticles to amplify an electrical signal. A sandwich assay may be carried out to recruit the gold nanoparticles to a magnetic bead by means of the target DNA. Silver can then be deposited. However, this method destroys the sample and requires many development steps. Reliability can also be an issue, although this is a very sensitive technique and could be suited to identifying more than one target strand simultaneously.[19]

Alternatively, a hybridisation indicator, or intercalator (DNA-mediated charge transport sensor), can be used, which contains the ferrocene molecule. The intercalator will be a molecule which, binds preferentially to double stranded DNA. This means that it will bind to the dsDNA formed from complementary strands of DNA (one strand containing the probe) without mismatches more often than to single strands. An example of this is the ferrocenylnaphthalene diimide threading intercalator from the group of Takenaka,[27,28] which binds much more quickly and strongly to a duplex than to a single strand (**Figure 7**).

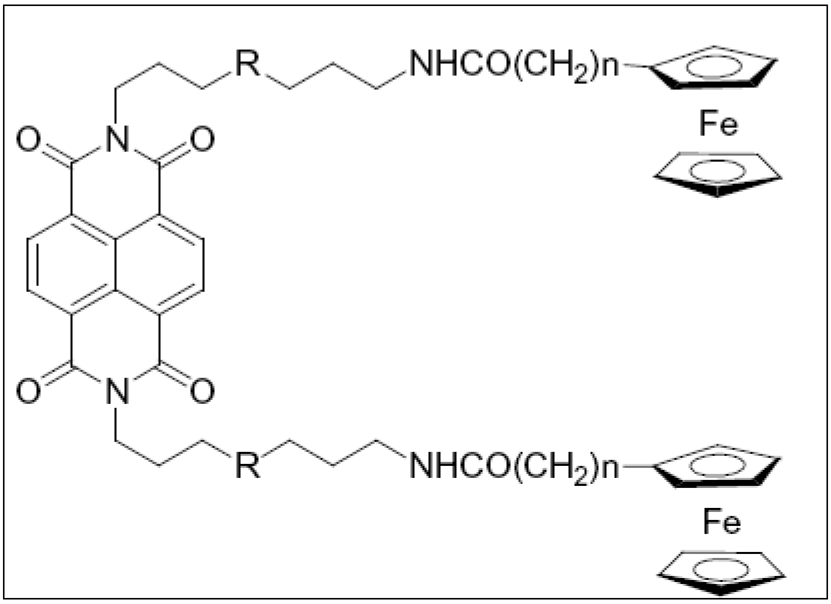

**Figure 7**. Example of a ferrocenylnaphthalene diimide intercalator[28].

The intercalators are sequence independent, and so can be used for a variety of oligonucleotides.

An electrochemical sensor consists of the sensing molecule, in this case ferrocene, incorporated into a known sequence of nucleotides to form a modified oligonucleotide. This strand is called a capture probe, and will interact with the strand of DNA that is being investigated (the target), as shown in **Figure 8**.

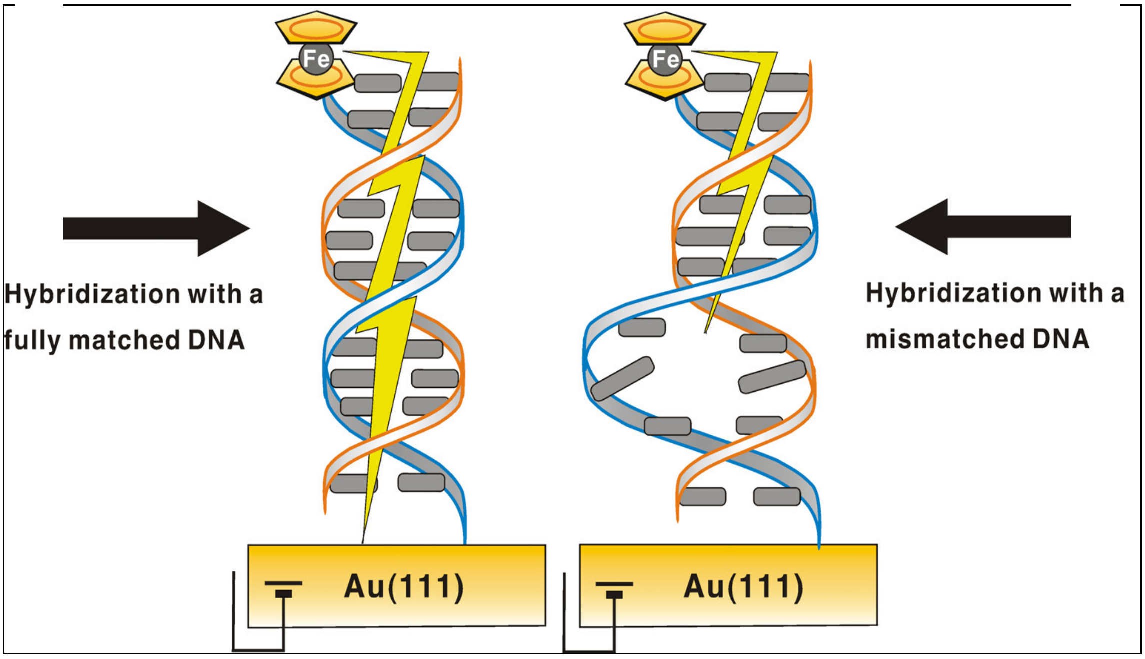

**Figure 8**. Schematic of electrochemical sensing of matched and mismatched ferrocene modified oligonucleotide sequences on a gold electrode surface[29].

The first reports of ferrocene incorporated into DNA were solution phase sensors. A ferrocene with a flexible linker was attached to the 5' end of an oligonucleotide, which in turn was hybridised to a complementary oligonucleotide strand, and this electrochemically active double stranded (ds) DNA was detected using HPLC-ECD.[14]

The ferrocene gene sensors currently being studied are generally supported on a surface, or hybridised to an unmodified oligonucleotide strand, which is attached to a solid surface. This allows the redox response of the ferrocene group to be easily detected.[25] The immobilisation process is important, as this will, to some extent, affect the reproducibility, sensitivity and selectivity of the sensor. As hybridisation of the probe and target occurs, both must be able to coil around each other, so the immobilisation of the probe must allow some configurational freedom for this to happen.[24]

The probe is generally immobilised onto a gold surface *via* a thiol linkage. The first group to immobilise the ferrocene on a gold surface created a ferrocenyl amidite, which was coupled to a thymidine modified controlled pore glass bead. This was followed by a 14 thymidine strand attached to the monomer and a thiol modified (on the 3' end) monomer, which was adsorbed onto a gold surface. Gold is a good surface for adsorption of monolayers, because it can be made atomically flat and adsorption of thiols onto gold surfaces is a very well researched process. Gold can also be used as an electrode to characterise the redox active DNA strands.[30]

The gold electrode surface can also be coated with a self assembled monolayer, consisting of the DNA capture probes, which allows electron transfer between the immobilised ferrocenes and the gold surface.[31] This can be seen in **Figure 7**.

All of the probes modelled on ferrocene are amperometric (they function by measuring the current associated with the oxidation of ferrocene at a constant potential)

## *1.6 Metal-based DNA Constructs*

As well as investigating the possibilities of ferrocene electrochemical sensors, we have become interested in the possibility of extending the theory of attaching a ferrocene moiety onto a strand of DNA, to fully incorporating one into the structure of the DNA helix. This would allow mismatch sensing to take place and could also facilitate construction of DNA nanoarchitectures or molecular electronic devices based on DNA. The redox properties of ferrocene point to the prospect of facilitating charge transport through DNA, potentially even making the theory of DNA based molecular wires possible.

There are already examples of metals incorporated into oligonucleotides, for a number of purposes. Metals have been covalently attached to oligonucleotides for energy transfer, electron transfer through DNA, construction of synthetic endonucleases and as nodes for the assembly of artificial DNA constructs.[32]

There are numerous examples of metals co-ordinating to base pairs, the first suitable ligand (palladium) being reported in 1999 by Tanaka and Shionoya for metal assisted base pairing.[33] More recently, mixed metal stacks have been demonstrated, with up to ten metal atoms per duplex.[32]

Mercury can bind to thymine and form T-Hg-T mismatches, with inter- and intrastrand crosslinks, sitting in the centre of the duplex (**Figure 9**). Initially, only three mercury ions could be inserted into the duplex. Other divalent ions that may form complexes include zinc, cobalt and nickel.[32]

As yet, there are no examples of a ferrocene molecule with suitably appended DNA bases being incorporated into the backbone of DNA, although attachment of ferrocenes to DNA is fairly common (as described in the previous section), with the closest examples seen in the work of Brisset.[34]

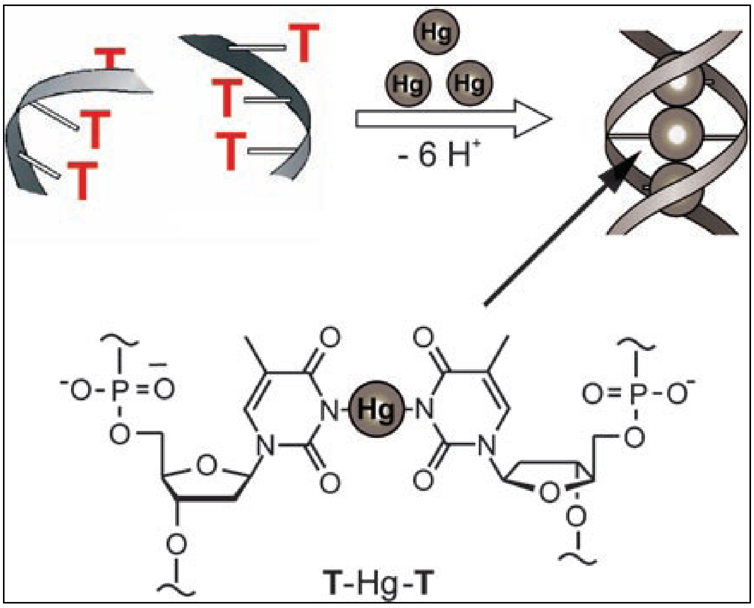

**Figure 9**. Example of a T-Hg-T base pair, with multiple mercury atoms incorporated into the duplex[32].

Work by Inouye and Takase[35] from 2001 shows the functionalisation of the two ferrocene Cp rings with diaminopyridine moieties, which can then bind non-covalently to a thymine based dinucleotide (**Figure 10**). The diamino pyridine moieties were coupled to the ferrocene *via* an iodine on each Cp. This work shows that such modified ferrocenes can be complexed to nucleotides in a fashion reminiscent of complementary base pairing in DNA.

It was noted in this publication that the distance between the ferrocene Cp rings was 0.33 nm. This is very similar to the distance between base pairs in a B DNA strand (as mentioned in section 1.3), giving the idea that a suitably functionlaised ferrocene might replace two DNA bases in the helix.

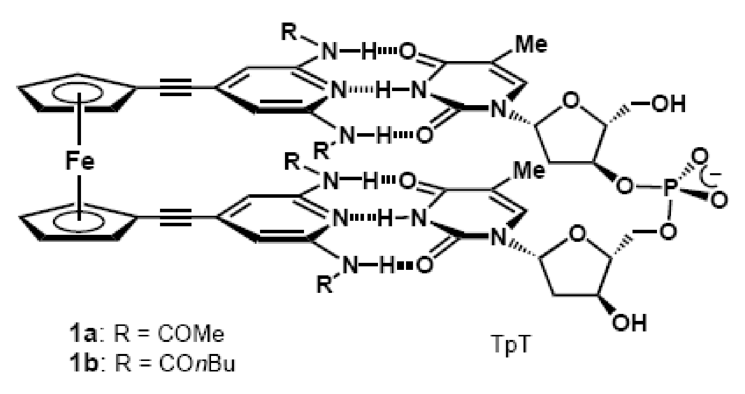

**Figure 10.** Ferrocenes with diaminopyridine moieties attached to a thymine based dinucleotide[35].

In addition, A. E. Beilstein and M. W. Grinstaff synthesised an oligonucleotide with ferrocene attached to uridine (the RNA replacement for T) or adenosine in the centre of the strand as tagging molecules,[36] showing the possibilities of incorporating ferrocene modified bases into DNA and Yu et al. attached a ferrocene to a deoxyribose sugar nucleotide (A and C) *via* a butoxy linkage.[37]

## *1.7 DNA synthesis*

Many of the examples given in this section use automated DNA synthesis techniques to incorporate the ferrocene derivatives that have been synthesised into an oligonucleotide strand. The most commonly used automated DNA synthesis technique is that using phosphoramidite chemistry, as developed by Caruthers and McBride.[38] This has replaced older methods, as it requires less particular conditions.

The generic DNA synthesis cycle is shown below (**Figure 11**).

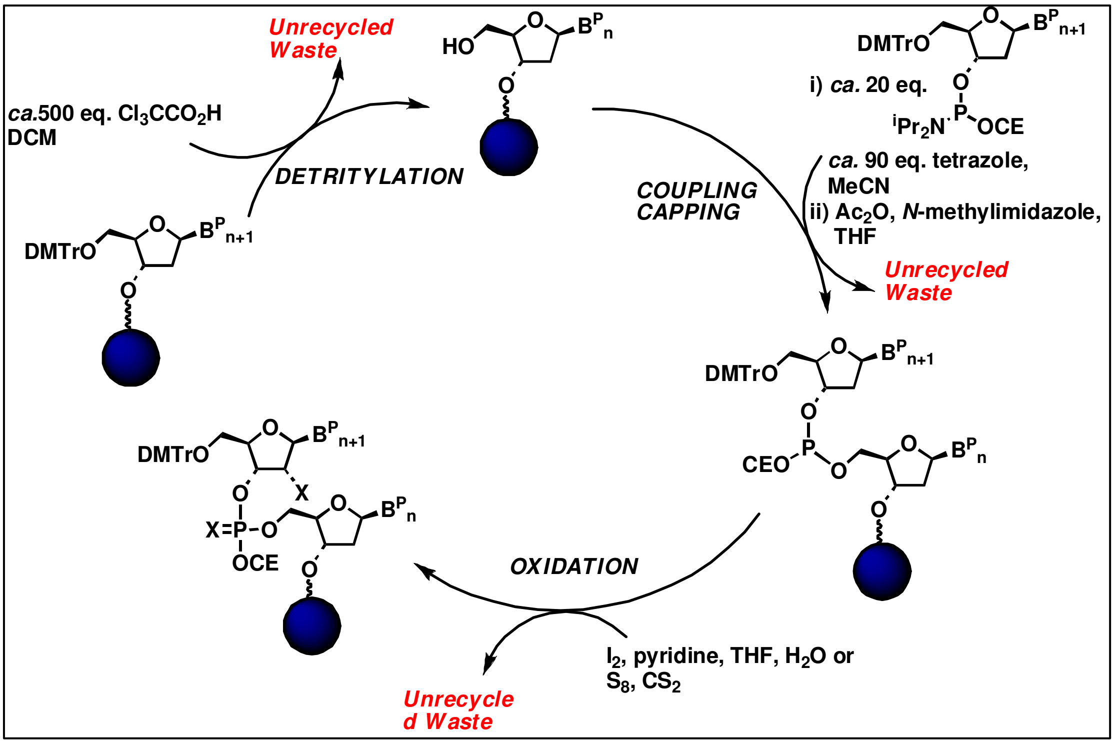

**Figure 11.** General DNA synthesis cycle. Image courtesy of Dr. J. Vyle, Chemistry Department, Queens University Belfast

This modern synthesis takes place on a solid support (controlled pore glass bead, CPG), proceeds in the 3’- 5’ direction and uses protected phosphoramidite base monomers (A, C, T, G) and tetrazole catalysis. The steps involved in the synthesis of an oligonucleotide are as follows:

Detritylation involves removing the dimethoxytrityl group from the 5’ OH of the first base, which is attached to the solid support.

During the coupling stage, the phosphate group of a nucleotide is activated by tetrazole, allowing it to couple (by condensation) to the detritylated 5’OH base and form a phosphite linkage, which is rather unstable. The tetrazole forms a tetrazolyl phosphoramidite intermediate with the monomer, which reacts with the hydroxyl group, forming a 5’-3’ linkage. The use of tetrazole increases the coupling efficiency to greater than 99%.

Capping prevents unreacted nucleotides (coupling failures) from reacting further, by attachment of an acetylating protecting group to the unreacted 5’ hydroxyl of the base. The column is then washed to remove any excess reagent before the next step in the cycle.

Oxidation stabilises the phosphite linkage between the two bases, converting it into a phosphate linkage. This is achieved using iodine as an oxidant. The iodine forms an adduct with the phosphite, which can then be decomposed by water, leaving the stable phosphate[39].

This sequence of steps is repeated to obtain an oligonucleotide strand of the desired length and succession of bases. In the case of the ferrocene phosphoramidite monomers, they were to be placed on the end of the sequence and in the middle. Once the desired oligonucleotide has been completed, it must be cleaved from the solid support using methylamine, and deprotected before being characterised by mass spectrometry.

At this point, it is still possible that there are undesirable by products, such as unreacted base monomers, present.

## *1.8 General and Specific Aims*

The overall aim of this research project was to develop a ferrocene derivative for incorporation into DNA *via* the established method of phosphoramidite solid state DNA synthesis and for use of the resulting redox-active oligomucleotides as sensors for mismatched DNA bases.

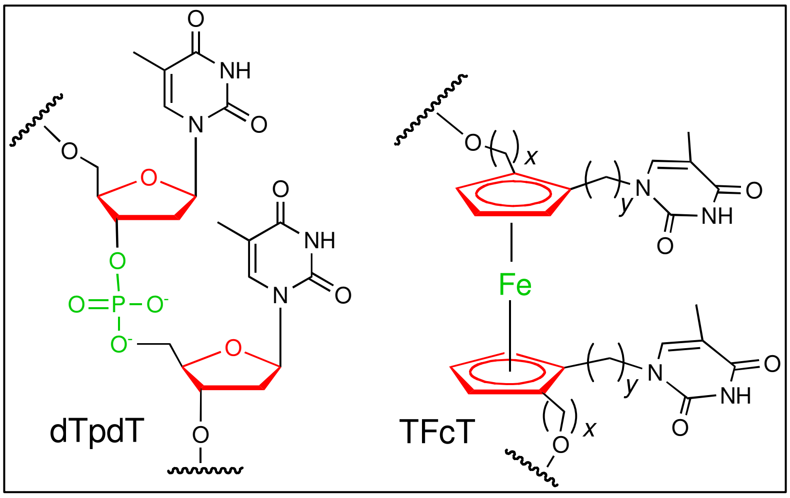

**Figure 12.** A standard TT dinucleotide (left) for comparison with the ferrocene nucleotide (right). Image courtesy of Dr. J. H. R. Tucker

Initially, the aim was to find a suitable synthetic route that would enable the functionalisation of ferrocene in such a way as to facilitate the synthesis of a metallocene nucleobase suitable for incorporation into DNA. It was thought that this could be achieved by creating a tetra-substituted ferrocene derivative (as shown in **Figure 12**), and one objective was to create a synthetic route towards this.

The spacing between the Cp rings is 3.3 Å; sufficiently similar to that of B-DNA for it to seem feasible that the stability of the double helix would not be compromised, assuming that the normal hydrogen bonds and stacking interactions can take place. **Figure 13** is a schematic of the DNA strands with the ferrocene nucleotide incorporated, showing the potential for a redox triggered hybridisation process.

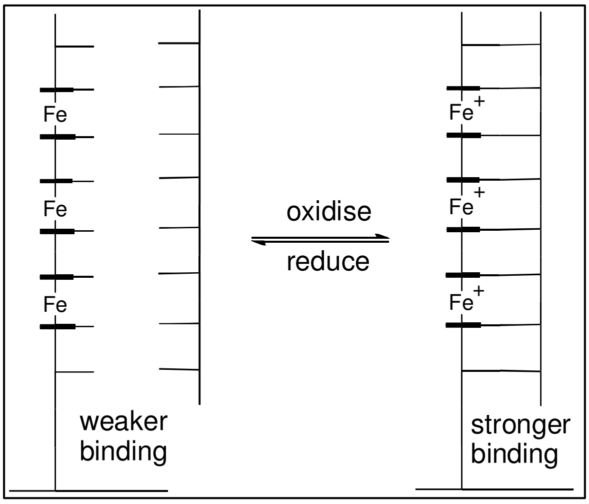

**Figure 13**. Schematic of the ferrocene nucleotide incorporated into the duplex and the redox-switching possibilities.

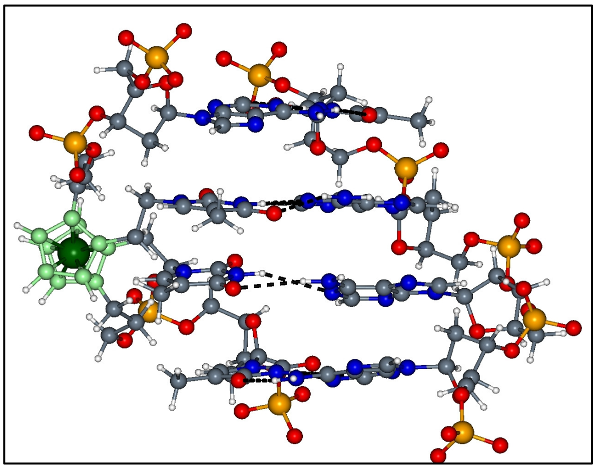

**Figure 14.** Computer modelled image of a ferrocene (pictured in green) nucleobase incorporated into DNA with functionalisation at the 1 and 2 positions. The sequence from the 5’ end is ATTT. x here is equal to three, and y is equal to one, with regards to figure 12. Image courtesy of Dr. T. Walsh, Department of Chemistry, Centre for Scientific Computing, University of Warwick

Preliminary computer modelling by Dr. Walsh at Warwick (**Figure 14**) has shown that the TFcT structure above where x = 3 and y = 2 (**Figure 12**) could be integrated into the B DNA duplex structure. The bases and the phosphate groups in the 1 and 2 positions of the rings would provide minimum disruption compared to attachment at the 1 and 3 positions of the cp ring.

As the tetra-substituted TFcTstructure is quite complex and would require many synthetic steps, it was thought prudent to also attempt to insert simpler disubstituted ferrocene derivatives into an oligonucleotide strand to see how feasible the concept was and to readily assess the effect of factors such as chain length on DNA stability. As mentioned earlier, some researchers have already successfully inserted ferrocene moieties into DNA strands, using alkyl chains as spacers between the ferrocene and the phosphate linkage [34]. This group did not, however introduce any molecules into the middle of the strand, which would be most useful for the proposed outcome of this project. This is in contrast to other groups, who have attached ferrocene onto (rather than into) DNA strands, as discussed previously (in section 1.5) with respect to electrochemical sensing of base mismatches.

The synthetic process was therefore set out to involve synthesis of disubstituted ferrocenes, which would be incorporated into the DNA backbone, and then eventually to move onto adding another two functional groups (i.e. DNA bases), giving a tetra-substituted ferrocene, which could be the desired ferrocene nucleotide, TFcT. The target disubstituted ferrocenes 1, 2 and 3 containing either amide or ether linkages are set out below in **Figure 15**.

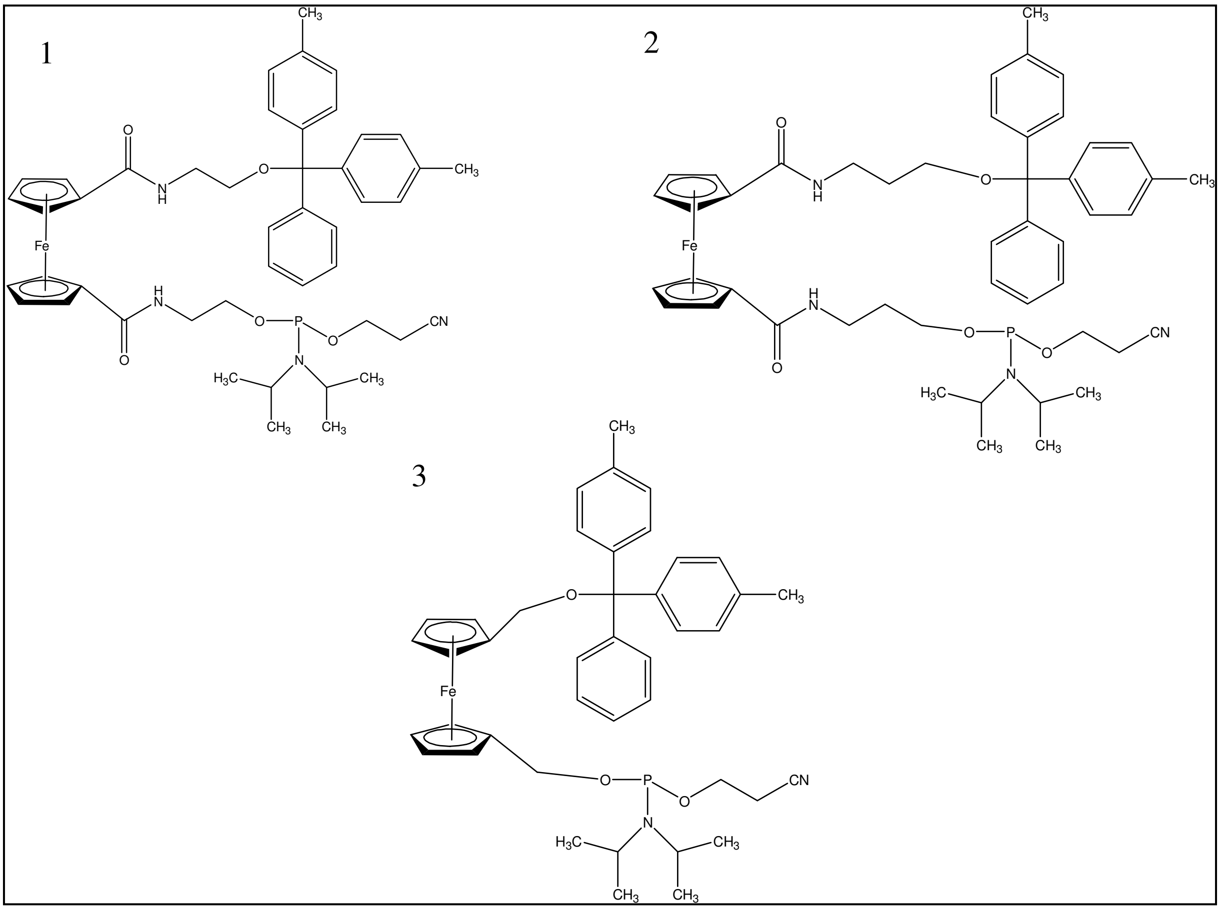

**Figure 15** Target compounds for the project. Compounds 1, 2 and 3 are simple phosphoramidites for incorporation into an oligonucleotide strand.

Amides have been previously used successfully in relation to DNA attachment with naphthyl[40] and pyrene[41] fluorescent detection compounds. There has also been research on ferrocene polyamide compounds as redox active DNA binding molecules[42] (**Figure 16**).

The amide functional group was chosen because it affords some rigidity with the CONH bond, but has the more flexible alkyl spacer between the amide bond and the alcohol. The amide chemistry is also already established and should not provide solubility problems under biological conditions.

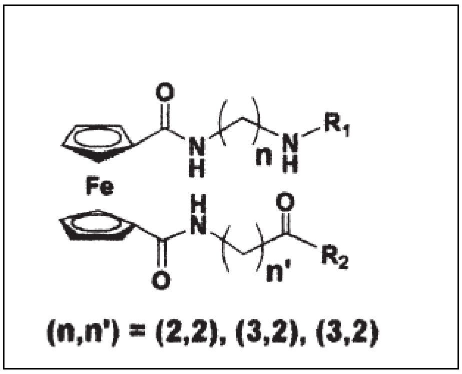

**Figure 16**. Structure of ferrocene polyamide compounds. Image from Mizuta, M., et al., Nucleic Acids Symposium Series, 2004. *48*, 237.

Although there is little flexibility in the system containing the hydroxy methyl group (compound 3), and the chain length would make it rather short to fit well into a DNA duplex without distorting it, the diol precursor is commercially available and it was thought this would provide a relatively easy synthetic route to DNA incorporation

The stability of the strands was to be tested to verify the feasibility of replacing a nucleotide unit with the ferrocene monomer described. Using different types of molecule would mean that there could be a comparison between a range of DNA strands of varying duplex stability, allowing the best functional groups and linker lengths to be identified and compared with modelling studies. Successful insertion of a ferrocene molecule would also provide some evidence that it would be possible to synthesise an oligonucleotide strand containing units such as TFcT. Once this was achieved, a complementary strand would then be synthesised and the non-covalent interactions and stability studied, as compared to the wild type strand (no modifications). The possibility of redox-switchable DNA could then also be investigated.

As described in subsequent chapters, the complete incorporation of a ferrocene-based nucleotide into DNA was not achieved within the time boundaries of this project, although it is the final aim, and further projects will hopefully go on to achieve this.

# 2 EXPERIMENTAL

Note: References listed as University of Birmingham Chemistry Department are experimental procedures taken from the undergraduate teaching labs.

## *2.1 Solvent and Reagent Pre-treatment*

All experiments were carried out using Schlenk Techniques and under nitrogen, with the exception of the oligonucleotide synthesis, where all reactions were completed under argon. Anhydrous diethyl ether and DCM were distilled over calcium hydride and the diethyl ether was stored over molecular sieves of 4Å diameter. THF was distilled over sodium and benzophenone.

## *2.2 Instrument Methods*

Mass Spectrometry was carried out using Electrospray techniques on a Waters/Micromass ZMD Electrospray Quadrupole Mass Spectrometer. MALDI-TOF was carried out using Bruker Biflex IV Maldi Time of Flight Mass Spectrometer. NMR was carried out using an AC300 ($^1$H) and AV300 Bruker NMR Spectrometer ($^{13}$C, $^{31}$P and 2D COSY). DNA synthesis was carried out using an Applied Biosystems Expedite 8900 Nucleic Acid Synthesis System, with desalting performed with Waters single use Sep-Pak® cartridges. HPLC was performed using a Merck Hitachi Interface D-7000, pump L-7100 with a La Chrom diode array detector L-7455. UV quantification was carried out using a Shimadzu UV3101PC spectrophotometer.

## *2.3 Synthesis of Monomers for Insertion into Oligonucleotides*

### 2.3.1 Preparation of 1'-((O-DMT-ethyl)amino)carbonyl)1’-((2-cyanoethyl-N,N-diisopropylphosphoramiditylethyl)amino)carbonyl) ferrocene[43] Compound 1)

#### 2.3.1.1 Preparation 1,1’-dilithioferrocene

A 3-necked round bottomed flask with a pressure equalised dropping funnel with stopper, septum and nitrogen inlet adaptor, was connected to a schlenk line. TMEDA (10 ml) and dry hexane (10 ml) were added *via* the septum and stirred under nitrogen. BuLi (42 ml) was then added and the mixture stirred to form the BuLi adduct (15 minutes, r.t.). Ferrocene (5.10 g, 0.027 mol) was dissolved in dry hexane (240 ml) and added to the reaction flask *via* the dropping funnel over a period of 30 minutes. The mixture was then left for 18 hours (under nitrogen) and the carboxylation carried out without purification

#### 2.3.1.2 Preparation of 1,1’- ferrocenedicarboxylic acid

Dry diethyl ether (500 ml) and an excess of crushed dry ice were placed into a beaker and the dilithioferrocene from 2.3.1.1 added with rapid stirring to precipitate a yellow/ orange salt of the dicarboxylic acid (precipitate). The mixture was warmed to room temperature and filtered, washed with dry diethyl ether (100 ml) and dried. The solid was dissolved in NaOH (600 ml, 0.1 M) and washed with dichloromethane (2 x 50 ml). The aqueous layer was retained and acidified by dropwise addition of HCl (conc. 37%, 9.61M) giving an orange precipitate. The solid was filtered and washed with water (100 ml), before being heated gently with toluene (100 ml) to remove any ferrocene monocarboxylic acid . The mixture was filtered once more, dried and weighed. The product weighed 4.34 g and a yield of 58% was recorded. The product was used without further purification in the next step.

#### 2.3.1.3 Preparation of 1,1'- bis(chlorocarbonyl)ferrocene[44]

The ferrocene dicarboxylic acid (4.34 g, 0.016 mol) and dry DCM (50 ml) were stirred under nitrogen and cooled to 0 °C using an ice bath. Excess Oxalyl Chloride (11.0 g) in dry DCM (20 ml) was added dropwise over a 30 minute period. DMF (4 drops) was added and the reaction warmed and left overnight under nitrogen. The solvent was removed *in vacuo* and the product weighed. The product was purified using a soxhlet extractor setup and pentane (200 ml) to extract. The solution was left for 18 hours under nitrogen. The resulting solution was evaporated to give a dark red, crystalline solid, weighed and stored under nitrogen in the freezer. The product weighed 1.37 g and a yield of 28% was recorded and $^1$H NMR analysis was carried out. The data corresponded to that in the literature and was used without further purification.

$^1$H NMR ( 300 MHz, $CDCl_3$): $\delta = 5.05- 4.65$ (m, 8H)

#### 2.3.1.4 Preparation of 1,1'-([bis(2-hydroxyethyl)amino]carbonyl)ferrocene[44]

1, 1’-bis-(chlorocarbonyl ferrocene (1.00 g, 3.22 mmol) was dissolved in dry THF (35 ml). 2-aminoethanol (0.01 mol, 0.6 ml) and triethylamine (0.01 mol, 5 ml) were added *via* a dropping funnel, and the reaction stirred overnight under nitrogen at room temperature. A brown precipitate was formed. The solution was filtered and washed with DCM (5 x 10 ml). The precipitate was washed with methanol (3 x 10 ml). The solvents were removed *in vacuo*. The precipitate was purified by column chromatography, using basic alumina and a gradient of methanol in DCM (5-8%). The product was removed as an orange band and the DCM was evaporated. $^1H$ and $^{13}C$ NMR analysis was carried out. The product weighed 0.52 g and a yield of 33% was recorded.

$^1H$ NMR ( 300 MHz, $CD_3CN$): $\delta = 5.48$ (s, 2H), 4.61-4.41 (m, 8H), 3.77-3.35 (m, 10H)

$^{13}C$ NMR ( 300 MHz, $CDCl_3$): $\delta = 75.79, 75.37, 74.95, 68.98, 68.21, 66.56, 61.39, 41.03$

#### 2.3.1.5 Preparation of 1-((O-DMT-ethyl) amino)carbonyl)1’-((hydroxyethyl) amino) carbonyl) ferrocene[34]

DMAP (0.34 mmol, 0.041g) and the amide from **2.3.1.4** (0.52 g, 1.44 mmol) were co-evaporated with dry THF (2 x 10ml). Dry THF (10 ml) was added to the residue along with DMT (0.65 g, 1.88 mmol,) and DIPEA (0.3 ml, 1.727 mmol) and the mixture stirred overnight under nitrogen at room temperature. The reaction was quenched with methanol (4 ml). The crude product was dissolved in DCM (25 ml) and the organic phase washed with saturated $NaHCO_3$ in water (5 x 25 ml) before being dried over magnesium sulphate. The solvents were removed and the product purified by silica gel chromatography (neutralised by 2% TEA) with a gradient of methanol in DCM (0.5 – 5%). The product was analysed by $^1$H and $^{13}$C NMR and Mass Spectrometry. The product weighed 0.29 g and a yield of 30% was recorded.

$^1$H NMR (300 MHz, $CD_3CN$): δ = 7.53-6.83 (m, 18H), 5.45 (s, 1H), 4.90 (1H), 4.59- 4.31 (s, 8H), 3.75 (s, 6H), 3.18-3.14 (t, J = 5.5 Hz, 2H), 1.29-1.24 (t, J = 7.0 Hz, 2H)

*m/z* (ES$^+$) calcd for $C_{37}H_{38}FeN_2O_6$ ([M + Na]$^+$) 662.55, found 662.30

### 2.3.1.6 Addition of phosphoramidite to 1-((O-DMT-ethyl) amino) carbonyl) 1’-((hydroxyethyl) amino) carbonyl) ferrocene[34] (Compound 1)

Dry DCM (1.4 ml) was added to the product from **2.3.1.5** (0.17 g, 0.28 mmol) and the mixture stirred under argon at room temperature. DIPEA (98 µl, 0.56 mmol) and 2-cyanoethyldiisopropylchlorophosphoramidite (77 µl, 0.31 mmol) were then added and the mixture was stirred for 15 minutes before being concentrated *in vacuo* without heating. The product was identified using $^{31}$P, $^1$H and and ES mass spectrometry, although it started to degrade before $^1$H and $^{13}$C NMR analysis could be carried out.

$^{31}$P NMR ( 300 MHz, $CD_3CN$): δ = 149.25, 147.83

*m/z* (ES$^+$) calcd for $C_{46}H_{55}FeN_4O_7P$ ([M + Na]$^+$) 862.78, found 862.3.

### 2.3.2 Preparation of (1-O-Dimethoxytrityl-2-hydroxy)diethylamine[34]

Procedure as for 2.3.1.5, replacing 1-((O-DMT-ethyl) amino) carbonyl) 1'-((hydroxyethyl) amino) carbonyl) ferrocene with diethanolamine.

Diethanolamine (0.14 g, 1.32 mmol) and DMAP (0.033 g, 0.26 mmol) were co-evaporated with dry THF (2 x 10 ml). Dry THF (10 ml) was added to the residue along with DMT (0.50g, 1.46 mmol) and DIPEA (1.32mmol, 0.23 ml) and the mixture stirred overnight under nitrogen at room temperature. The reaction was quenched with methanol (4 ml). The crude product was dissolved in DCM (25 ml) and the organic phase washed with saturated $NaHCO_3$ in water (5 x 25 ml) before being dried over magnesium sulphate. The solvents were removed and the product purified by silica gel chromatography (neutralised by 2% TEA) with a gradient of methanol in DCM (0.5 – 5 %). The crude product weighed 0.76 g and a yield of greater than 100% was recorded. The product was analysed by $^1H$ NMR and Mass Spectrometry.

$^1H$ NMR (300 MHz, $CD_3CN$): $\delta =$ 7.41-6.83 (m, 13H), 3.76-3.75 (d, J = 4.8 Hz, 4H), 3.53-3.49 (t, J = 5.5 Hz, 1H), 3.35-3.33 (m, 1H), 3.13-3.10 (t, J = 5.5 Hz), 2.96 (s, 1H), 2.81-2.74 (m, 1H), 2.67-2.63 (t, J = 5.5 Hz, 1H)

*m/z* ($ES^+$) calcd for $C_{25}H_{29}O_4$ ($[M + Na]^+$) 407.50, found 407.20

## 2.3.3 Preparation of 1-(O-Dimethoxytritylmethyl)-1'-(O-(2-cyanoethyl-N,N-diisopropylphosphoramidityl)methyl)ferrocene (Compound 3)

### 2.3.3.1 Preparation of 1-(O-Dimethoxytritylmethyl)-1'-(hydroxymethyl) ferrocene[34]

The procedure followed was the same as for **2.3.1.5**, but with 1, 1'- ferrocenedimethanol (0.16 g, 0.66 mmol). Column chromatography could not separate the mono- and bis- products sufficiently. A second column was prepared with a 3:1 mix of hexane to ethyl acetate respectively. The product weighed 0.28 g and a yield of 77% was obtained. The fractions were analysed by $^1$H NMR and Mass Spectrometry analysis.

$^1$H NMR (300 MHz, $CD_3CN$): $\delta =$ 7.46- 6.83 (m, 13H), 4.13-3.97 (m, 12H), 3.80 (s, 1H) 3.74 (s, 6H)

*m/z* (ES$^+$) calcd for $C_{33}H_{32}FeO_2$ ([M + Na]$^+$) 548.45, found 548.2

### 2.3.3.2 Preparation of 1-(O-Dimethoxytritylmethyl)-1'-(O-(2-cyanoethyl-N,N-diisopropylphosphoramidityl)methyl)ferrocene[34]

The procedure followed was as for **2.3.1.6**, using the product from **2.3.3.1** (0.28 g, 0.50 mmol), DIPEA (130 µl, 1.01 mmol) and 2-cyanoethyldiisopropylchlorophosphoramidite (187 µl, 0.55 mmol). The product was identified by $^{31}$P NMR spectroscopy.

$^{31}$P NMR ( 300 MHz, $CDCl_3$): $\delta$ = 149.19-148.26

## *2.4 Synthesis of oligonucleotide strands with ferrocenes from section 2.3 incorporated*

**Note:** Procedures for synthesis of DNA strands supplied by Dr. J. Vyle, Chemistry Department, Queens University Belfast

### 2.4.1 Addition of phosphoramidite to ferrocenes

#### 2.4.1.1 Preparation of 1-(O-dimethoxytritylmethyl),1'-(O-(2-cyanoethyl-N,N-diisopropylphosphoramidityl)methyl)ferrocene[34] (Compound 3)

1-(O-Dimethoxytritylmethyl)-1'-(hydroxymethyl) ferrocene (0.58 g, 1.067 mmol) was dried under vacuum overnight. DCM (13 ml) was then added and a sample taken for TLC. THF (2.3 ml) and DIPEA (0.76 ml, 4.38 mmol) were added to the solution, along with the phosphoramidite (0.26 ml, 1.17 mmol). The mixture was stirred under argon for 30 minutes, whereupon a sample was taken to run a TLC plate. TLCs were run in a mixture of DCM (3ml), Hexane (14 ml), ethyl acetate (2 ml) and TEA (1 ml) on silica gel plates. When the reaction was judged to have reached completion, the solution was quenched with benzyl polymer beads (0.68 g, 0.21 mmol), which had been dried overnight under vacuum. The mixture was then left stirring for 1 hour. After this time the solution was poured into ethyl acetate (30 ml, washed in $Na_2CO_3$), washed with $Na_2CO_3$ (2 x 10 ml) and dried over sodium sulphate before being reduced *in vacuo*. A mixture of dry DCM and dry hexane (30:70, 20 ml) and activated basic alumina were added to the residue and the resulting mixture stirred for 30 minutes, then filtered through activated basic alumina and washed with the dry DCM/

hexane solution. The filtrate solution was reduced *in vacuo,* and then co-evaporated twice with dry acetonitrile (10 ml) before being dried on a high vacuum line. A $^{31}$P NMR was used to identify the desired product was present and 0.29 g of product was obtained, with a yield of 39% recorded

$^{31}$P NMR ( 300 MHz, $D_2O$): $\delta = 150.79- 149.45$

### 2.4.1.2 Preparation of 1-((((O-dimethoxytritylpropyl)amino) carbonyl))-1'-((O-(2-cyanoethyl-N, N-diisopropylphosphoramiditylpropyl)amino) carbonyl) ferrocene[34] (Compound 2)

This molecule was not prepared for this project, but as it was of the same family of compounds, was used for incorporation into an oligonucleotide sequence. It was synthesised at the University of Birmingham by Jean-Louis Duprey.

The procedure was the same as for **2.4.1.1**, replacing the 1-(O-Dimethoxytritylmethyl)-1' (hydroxymethyl)ferrocene with 1-((O-Dimethoxytritylpropyl)amino)carbonyl)1'-((hydroxypropyl)amino)carbonyl)ferrocene (0.50 g, 0.74 mmol). DIPEA (0.53 ml, 3.03 mmol) and the phosphoramidite (0.18 ml, 0.81 mmol) were added. The filtration through activated basic alumina was omitted for this compound. The solution was instead washed with saturated NaCl (20 ml), after being poured into ethyl acetate and washed with $Na_2CO_3$. To facilitate the drying of the compound, a dry DCM/hexane (1:1) mixture was filtered through activated basic alumina and added to the residue at the co-evaporation stage. A $^{31}$P NMR was used for identification of the product. The product weighed 0.65 g and a yield of 100% was recorded.

$^{31}P$ NMR ( 300 MHz, $D_2O$): $\delta = 150.420- 149.077$

### 2.4.2 Attempted Synthesis of oligonucleotide sequences

Synthesis was carried out using a DNA synthesiser for the main strands at Queens University, Belfast.

The ferrocene was attached at the relevant position in the sequence *via* an off-column coupling using BTT. The ferrocene phosphoramidite compounds were dissolved in anhydrous acetonitrile. After the oligonucleotide synthesis (DMT on), the CPG support was treated with methylamine (50% aqueous) for 1 hour at 65°C, the supernatant recovered and evaporated. The oligonucleotide was then purified by reverse phase HPLC and detritylated. The oligonucleotides were eluted using a linear gradient of triethylammonium acetate buffer (5-40%) in aqueous triethylammonium acetate (100 μM, pH 6.5). The procedure for making buffer A is detailed in **appendix 1**, with details on desalting for mass spectrometry, UV quantification and $T_m$ studies. The main peaks were eluted at around 25 minutes and characterisation by MALDI-ToF mass spectrometry carried out, showing that the synthesis was unsuccessful.

## *2.5 Synthesis of Nucleobase Precursor*

### 2.5.1 Protection of 1,1'- ferrocenedimethanol with TBDMSCl[45]

1,1'-ferrocenedimethanol (1.00g, 4.06 mmol) was placed into a flask fitted with a reflux condenser along with TBDMSCl (1.40 g, 2.4 equivalents) and DMF (2 ml), with imidazole (0.69 g, 2.5 equivalents) as a catalyst. The mixture was left stirring for 48 hours at 40°C and then washed with sodium bicarbonate (2 x 25 ml), distilled water (2 x 25 ml) and brine (2 x 25 ml). The product was then dried over $MgSO_4$. A brown oil was obtained, which solidified. The product weighed 1.30 g and a yield of 67% was recorded. TLCs were carried out to determine whether the reaction had reached completion and $^1H$ and $^{13}C$ NMR and Mass Spectrometry allowed identification of the product.

$^1H$ NMR ( 300 MHz, $CD_3CN$): $\delta$ = 4.44 (s, 4H), 4.15-4.08 (m, 8H), 0.89 (s, 18 H), 0.06 (s, 12H)

$^{13}C$ NMR (300 MHz, $CD_3CN$): $\delta$ = 68.77, 68.32, 61.10, 25.29, -5.96

*m/z* ($ES^+$) calcd for $C_{24}H_{42}FeO_2Si_2$ ($[M + Na]^+$) 474.61, found 474.4

### 2.5.2 Attempted Lithiation of protected ferrocene

#### 2.5.2.1 Method 1[43]

The same method as described in section **2.3.1** was utilised, but using the TBDMS protected 1,1'-ferrocenedimethanol (0.25 g, 0.53 mmol). The reaction mixture dried out overnight and did not lithiate successfully, as evidence by $^1H$ NMR and mass spectrometry. The reaction was repeated with more solvent to prevent the reaction drying out. Lithiation did not occur, although a fine suspension was seen after the addition of the dry ice. $^1$H NMR and mass spectroscopy spectra revealed this to be non-ferrocene based.

### 2.5.2.2 Method 2[46]

TMEDA (5.3 equivalents, 0.17 ml) and n-BuLi (0.70 ml, 5.3 equivalents) in dry hexane (10 ml) were stirred at r.t. for ten minutes with the exclusion of moisture under nitrogen. The protected ferrocene synthesised using the procedure laid out in section **2.5.1** (0.1 g, 0.21 mmol) was added and the mixture stirred for 6 hours. The solution was expected to become darker as the reaction progressed, but this did not happen. After 6 hours, the mixture was cooled and chlorodiphenylphosphine (0.09 ml, 2.53 equivalents) added and refluxed overnight at 69° C. After this time, the reaction was quenched with distilled water (15 ml) and the organic layer separated and washed with more distilled water, before being dried over magnesium sulphate. The solvents were removed and samples analysed by $^1$H and $^{31}$P NMR and ES Mass Spectrometry analysis. The desired product was not present in the sample.

### 2.5.2.3 Method 3[47]

Protected ferrocene synthesised using the procedure laid out in section **2.5.1** (0.57g, 1.00 mmol,) was dissolved in dry hexane and n-BuLi (1.95 ml, 3.12 mmol) added dropwise. This was stirred for 2.5 hours before an excess of dry ice and diethyl ether were added as in 1.1. The product was analysed with ES Mass Spectrometry and $^1$H NMR, but the desired compound was not present.

## 2.5.3 Attempted Preparation of 1,1'-dimethylaminoethylferrocene

### 2.5.3.1 Attempted Acetylation of ferrocene (Method 1)[48]

Ferrocene (0.500 g, 1.56 mmol) was dissolved in DCM (40 ml) and added dropwise over a 1 hour period to aluminium chloride (0.63 g, 4.68 mmol) and acetyl chloride (0.34 ml, 4.7 mmol) in DCM (40 ml). The solution was then refluxed for 20 hours. After this time, the reaction was hydrolysed with deionised water and the aqueous hydrolysate extracted into chloroform. The majority of the solvent was removed *in vacuo* and the solution left to form crystals. The crystals were dissolved in toluene, washed with sodium bicarbonate, dried over sodium sulphate and filtered with the remaining solvent being removed *in vacuo*.

A column was run on silica, with ethyl acetate and hexane (1:1). This literature preparation was unsuccessful, resulting in formation of mainly mono- acetyl ferrocene.

### 2.5.3.1.2 Acetylation of ferrocene (Method 2)[43]

Ferrocene (3 g, 0.016 mol) was dissolved in dry DCM (15 ml) and added dropwise over a period of 15 minutes to a solution of aluminium chloride (5.5 g, 0.04 mol) and acetyl chloride (1.167 ml, 0.015 mol) in dry DCM (20 ml). The solution was then stirred for 2 hours at r.t. The mixture was then cooled to 0° C and ice added. The resulting mixture was then filtered and the solid residue washed until colourless with DCM (5 x 10 ml). The red solution was transferred to a separating funnel and separated from the aqueous phase, which was then washed with DCM (3 x 10 ml). The combined red phases were then dried over anhydrous magnesium sulphate and reduced to a volume of approximately 20 ml, after which point cyclohexane (10 ml) was added. The solution was concentrated again to 20 ml and allowed to cool slowly. The crystals formed were collected by filtration and air-dried. The product weighed 4.60 g and a yield of 74% was recorded. $^1$H and $^{13}$C NMR and ES Mass Spectrometry spectra were obtained to confirm the product.

$^1$H NMR (300 MHz, $CDCl_3$) δ: 4.77 (s, 4H), 4.51 (s, 4H), 2.35 (s, 6H)

$^{13}$C NMR ($CDCl_3$ , 300 MHz) δ = 198.88, 79.25, 71.82, 69.14, 25.17

*m/z* ($ES^+$) calcd for $C_{14}H_{14}FeO_2$ ($[M + Na]^+$) 270.1, found 270.1

Literature values[49]:$^1$H NMR (200 MHz, $CDCl_3$): δ = 4.73-4.72 (m, 4H), 4.47-4.46 (m, 4H). 2.31 (s, 6H)

$^{13}$C NMR ($CD_3CN$, 50 MHz) δ = 201.10, 80.55, 73.49, 70.85, 27.54.

#### 2.5.3.2.1 Reduction of 1,1’-diacetylferrocene giving 1,1'-Bis(a-hydroxyethyl)ferrocene (Method 1)[50]

Lithium aluminium hydride (0.13 g, 7.51 mmol) in anhydrous ether (5ml) was added dropwise to a solution of diacetylferrocene (0.91 g, 3.38 mmol) dissolved in anhydrous ether (10 ml) and heated for 2 hours under reflux. Excess lithium aluminium hydride was destroyed with ethyl acetate (5ml) and ammonium chloride (0.90 g) in water was added. The mixture was stirred at 0 °C for 30 minutes before being filtered and separated. The organic layer was washed with water before being dried and filtered. The solvent was removed *in vacuo.* The product weighed 0.47 g and a yield of 51% was recorded. Analysis by mass spectrometry and $^1$H and $^{13}$C NMR confirmed reduction was successful.

$^1$H NMR (300 MHz, $CD_3CN$) δ: 4.56-4.54 (m, 2H), 4.22-4.08 (m, 8H), 3.89 (s, 2H), 1.34-1.32 (dd, 6H)

$^{13}C$ NMR (300 MHz, $CDCl_3$): δ = 95.54, 80.98, 79.68, 77.02, 70.32, 70.32, 69.27, 68.23, 67.81, 66.98, 66.65, 66.59, 66.57, 66.33, 63.86, 26.07, 25.62, 25.70, 21.23

*m/z* ($ES^+$) calcd for $C_{14}H_{18}FeO_2$ ($[M + Na]^+$) 274.14, found 274.0

### 2.5.3.2.2. Method 2 [51]

Anhydrous ether (25 ml) and 1,1'-diacetylferrocene (0.75 g, 2.79 mmol) were heated to slow reflux with stirring. Lithium aluminium hydride (0.11 g, 2.79 mmol) in anhydrous ether (10 ml) was added dropwise and the solution was then refluxed for 15 minutes. Excess lithium aluminium hydride was destroyed with ethyl acetate, ice and water and the solution cooled to 0 ºC. An aqueous solution of ammonium chloride (1.10g in 10 ml water) was added and the reaction mixture stirred for 15 minutes, after which any solid lithium aluminium hydride by-products were removed by filtration. The aqueous phase of the solution was removed, and the organic phase washed with water (2 x 15 ml), before being dried over magnesium sulphate, filtered and the solvent removed. Orange/yellow oil was obtained, which was suitable for use without further purification. The product weighed 0.78 g and a yield of 100% was recorded.

$^1H$ and $^{13}C$ NMR and ES Mass Spectrometry spectra confirmed the presence of the required product, which consisted of a mixture of diastereomers.

$^1H$ NMR (300 MHz, $CDCl_3$): δ= 4.93 (s, 2H), 4.70-4.58 (m, 2H), 4.27-4.06 (m, 8H), 1.46-1.37 (m, 6H)

$^{13}C$ NMR (300 MHz, $CDCl_3$): δ = 94.18, 76.49, 76.06, 75.64, 67.29, 66.89, 66.77, 66.61, 66.52, 66.36, 65.52, 65.14, 64.94, 64.51, 64.19, 24.62, 24.24

*m/z* ($ES^+$) calcd for $C_{14}H_{18}FeO_2$ ($[M + Na]^+$) 274.14, found 274.0

Literature values for mixture of isomers[49]: $^1$H NMR (300 MHz, CDCl$_3$): δ= 5.15 (s, 1H) 5.12 (s, 1H), 4.64 (q, J =6 .2 Hz, 1H), 4.60 (q, J = 6.2 Hz, 1H), 4.25-4.11 (m, 8H), 1.39- 1.36 (dd, J = 6.6 Hz, 6H)

$^{13}$C NMR (300 MHz, CDCl$_3$): δ = 95.20, 67.74, 67.57, 67.48, 67.32, 66.52, 66.13, 65.92, 65.50, 65.40, 65.06, 25.56, 25.16

### 2.5.3.2 Attempted Preparation of 1,1'-bis(α-acetoxyethyl)ferrocene[52]

1,1'-Bis(1-hydroxyethyl)ferrocene (0.31g, 1.12 mmol) was dissolved in dry toluene (20 ml) with glacial acetic acid (0.23 ml) and set up with a water separator (Dean and Stark trap) and reflux condenser. The solution was refluxed and more toluene (40 ml) added to assist the water separation. The solution was refluxed until no more water separated, then cooled and the solvent removed *in vacuo*. Samples were analysed using $^1$H and $^{13}$C, HSQC and COSY NMR, Mass Spectrometry and IR. The product was not conclusively found to be present.

### 2.5.3.3.2 Method 2[49]

Acetic anhydride (0.5. ml, 5.29 mmol) was added to 1,1'-bis(1-hydroxyethyl)ferrocene (0.20 g, 0.74 mmol) in pyridine (7 ml) and the solution stirred at r. t. for 12 hours. After this time, the removal of the volatile matter under vacuum was attempted unsuccessfully. The mixture was then separated with water and extracted into DCM, dried over magnesium sulphate and the solvent was removed *in vacuo.* Purification was carried out by columning on silica gel neutralised with triethylamine (2%) and a 1:1 mix of ethyl acetate and hexane eluent. However, the product was not present on analysis of the fractions by $^1$H NMR and ES Mass Spectrometry.

### 2.5.3.3.3 Method 3[53]

Procedure as above, but the solution was quenched with water and extracted into diethyl ether, before being dried over magnesium sulphate and the solvent removed *in vacuo.* The compound obtained was analysed using $^1$H NMR and ES Mass Spectrometry and found not to be the desired product.

### 2.5.3.3.4 Preparation of 1,1'-bis(α-acetoxyethyl)ferrocene (Method 4)[49]

Procedure was repeated as above, but the reaction was poured into saturated aqueous $NaHCO_3$ (20 ml) and extracted into diethyl ether (40ml) after stirring. The organic phase was washed with water (20 ml) and brine (20 ml), then dried over magnesium sulphate and solvent evaporated to produce a yellow oil, which was used in the next step with no further purification. Product obtained was a mixture of diastereomers weighing 0.28 g and the yield was 43%. The product was analysed by ES Mass Spectrometry and $^1$H and $^{13}$C NMR.

$^1$H NMR (300 MHz, $CDCl_3$): $\delta$ = 5.80-5.78 (q, J = 6.3 Hz, 2H), 4.23-4.12 (m, 8H) 2.03 (s, 6H), 1.541- 1.510 (m, 6H)

$^{13}$C NMR (300MHz, $CDCl_3$) : $\delta$ = 172.66, 172.63, 91.01, 90.93, 79.91, 79.48, 79.06, 75.84, 73.23, 71.46, 71.20, 71.10, 70.80, 68.94, 68.91, 23.63, 22.57, 22.37

*m/z* ($ES^+$) calcd for $C_{18}H_{22}FeO_4$ ($[M + Na]^+$) 358.21, found 358.2

Literature values (after stereoselective reduction)[49]: $^1$H NMR (300 MHz, $CDCl_3$): $\delta$ = 5.72 (q, J = 6 Hz, 2H), 4.17-4.14 (m, 2H), 4.10-4.08 (m, 2H), 4.06-4.04 (m, 4H), 1.95 (s, 6H), 1.46-1.45 (m, 6H)

$^{13}$C NMR (75MHz, $CDCl_3$) : $\delta$ = 169.86, 169.84, 88.50, 88.43, 68.80, 68.56, 68.46, 68.14, 68.11, 66.34, 66.32, 20.97, 20.04, 19.84

### 2.5.3.3 Attempted Preparation of 1,1’-bis(α- N,N-dimethylaminoethyl)ferrocene (Compound 4)[49]

The acetate from **2.5.3.3** (0.28g, 0.79 mmol) was dissolved in methanol (10 ml) and an excess of dimethylamine (2 ml, 40% in water) and water (2 ml) added. The mixture was stirred overnight and then poured into saturated aqueous ammonium chloride (50 ml) and extracted into ether (100 ml). The organic layer was washed with water (2 x 50 ml) and brine (50 ml), then dried over magnesium sulphate and the solvent removed. Purification was carried out by column chromatography on silica, with diethyl ether and triethylamine (95:5)

No product was present, as determined by ES Mass Spectrometry and $^{1}$H NMR

The experiment was repeated, but left for 48 hours and then heated for three hours. No product was present, as ascertained by $^{1}$H NMR and ES mass spectrometry.

## 2.5.4 Preparation of 1,1’-bis-(dimethylaminomethyl)ferrocene[49]

### 2.5.4.1 Preparation of 1,1’-bis-(acetoxymethyl)ferrocene

Procedure as **2.5.3.3,** replacing the 1,1’-bis-(1-hydroxyethyl)ferrocene with 1,1’-ferrocenedimethanol (0.719 g. 2.92 mmol). The product was analysed by ES Mass Spectrometry and $^{1}$H and $^{13}$C NMR.

$^{1}$H NMR (300 MHz, $CDCl_3$): $\delta = 4.77$ (s, 4H), 4.19-4.106 (m, 8H), 1.89 (s, 6H)

$^{13}$C NMR (300 MHz, $CDCl_3$): $\delta = 79.24$, 78.82, 78.39, 71.86, 71.20, 64.30, 22.76

*m/z* ($ES^+$) calcd for $C_{16}H_{18}FeO_4$ ($[M + Na]^+$) 330.16, found 330.0

### 2.5.4.2 Attempted Preparation of 1,1'-bis-(dimethylaminomethyl)ferrocene

Procedure as **2.5.3.4**, using 1,1'-bis-(acetoxymethyl)ferrocene (0.3597 g, 1.09 mmol)

No product was present, as determined by ES Mass Spectrometry and $^1$H NMR

Procedure as above, but THF (10 ml) used instead of methanol.

No product was present, as determined by ES Mass Spectrometry and $^1$H NMR

# 3 RESULTS AND DISCUSSION

## *3.1 Phosphoramidite Monomer Synthesis*

One of the first things to consider for synthesis of a ferrocene nucleotide, such as TFcT, was whether the ferrocene could be feasibly integrated into the DNA strand. This had been successfully achieved a number of times where the ferrocene was attached to the end of a strand, and also in the centre of a strand, although less often, as described in section 1[12, 23-27, 34, 35]. Amides (with only one of the Cp rings functionalised and with a 6 carbon chain)[54] and alkyl chain ferrocenes[34] (albeit with longer chains than synthesised here) have been previously inserted into DNA, with the latter being the closest example of incorporation into the backbone. This suggested that the molecules chosen here could also be successfully integrated.

The simpler phosphoramidite targets were two different ferrocene compounds, one with an amide and an hydroxy ethyl group on each Cp (compound 1), and one with a hydroxy methyl group (no amide, compound 3) on each Cp ring. A third molecule, an amide identical to compound 1, but with an extra carbon in the alkyl chain, had been previously synthesised in the group (compound 2,[55] all shown below in **Figure 17**). This molecule was deemed relevant to the project, and was also included for subsequent oligonucleotide synthesis.

One of the alcohol functional groups was coupled to a phosphoramidite molecule (**Figure 17**), which acts in a similar fashion to the phosphate on a standard nucleotide and allows for incorporation into the oligonucleotide sequence, using automated DNA synthesis (as described in **section 1.7**), while the other alcohol was protected with a dimethoxytrityl chloride (DMT) group. This is a common protecting group in DNA synthesis, and one that is used for standard oligonucleotide strand synthesis.

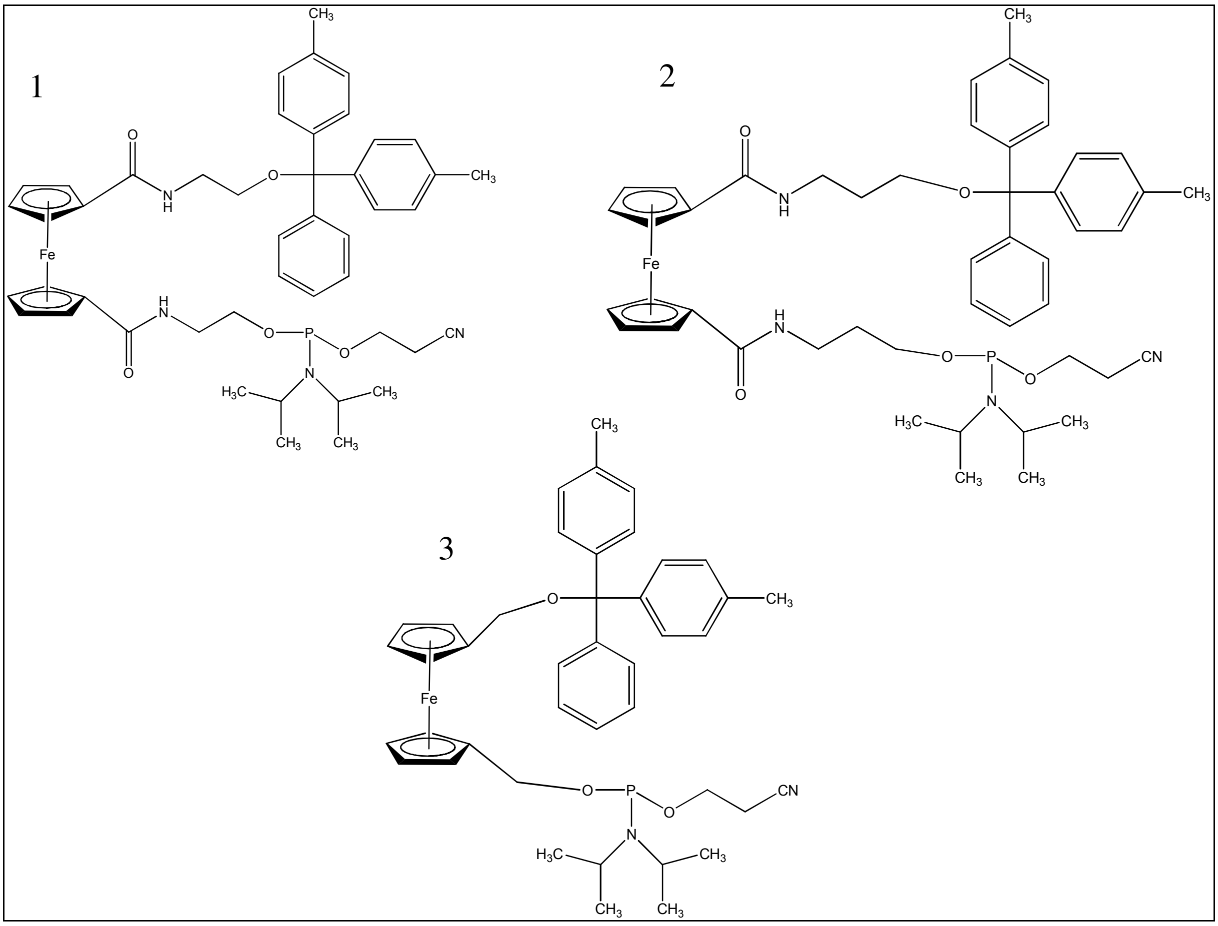

**Figure 17.** Phosphoramidite molecules for insertion into oligonucleotides. Compounds 2 and 3 were inserted into sequences, but 1 was not, although the phosphoramidite was synthesised.

Although time constraints meant that an oligonucleotide strand of the amide with a linker of two carbons (compound 1, above) could not be synthesised, the other two molecules were incorporated into DNA using a combination of off and on column coupling with a DNA synthesiser. These were purified by HPLC and the melting temperature of the duplexes ($T_m$) was measured.

## 3.2 *Synthesis of phosphoramidite monomers for insertion into an oligonucleotide strand.*

### 3.2.1 Compound 1

Initially, 1,1'-([bis(2-hydroxyethyl)amino]carbonyl)ferrocene was synthesised from ferrocene, using a well-established procedure. The ferrocene was lithiated using n-butyllithium and dry ice used to turn it into 1,1'-ferrocenedicarboxylic acid. This compound was then converted to 1,1'-bis(chlorocarbonyl)ferrocene using oxalylchloride. This synthetic scheme is shown in **Figure 18**.

A DMT moiety was then attached, *via* one of the alcohols, as a protecting group for the DNA synthesis, using the procedures laid out by Navarro et al.[34]

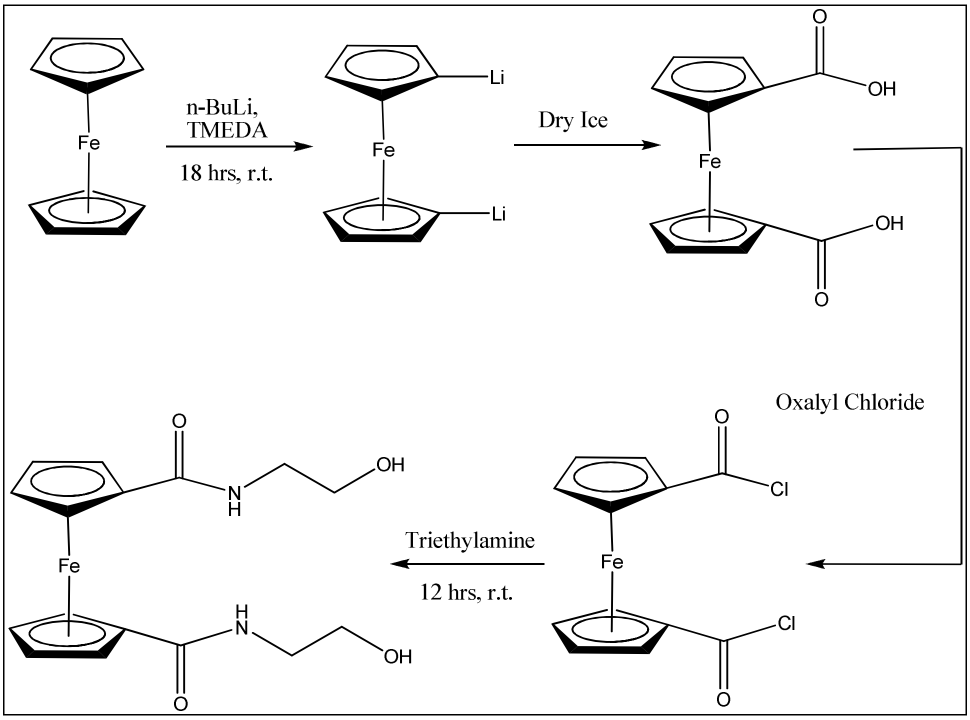

**Figure 18**. Schematic for the synthesis of the bisamide (shown here for n=2, compound 1)

The phosphoramidite was added to the ethyl amide to test whether the reaction would work. A positive result was identified by $^{31}$P NMR, giving compound 1 and proving that it would be possible to attach a phosphoramidite to these molecules and use them for DNA synthesis. The NMR was used for identifying whether the product was present, with a peak at approximately 150 ppm. This was not used to confirm the purity of the product, as it degraded very quickly and degradation products were prominent on the spectrum. This was true for all the phosphoramidites synthesised. The compound was synthesised in quantitative yield and the intermediate steps were characterised, using a range of spectroscopic techniques.

This particular molecule was not incorporated into an oligonucleotide due to time restraints, and it is expected that it will be attempted in future work.

### 3.2.2 Compound 2 (n3 on experimental data)

The phosphoramidite was added to this molecule and the product identified by $^{31}$P NMR, which showed a peak at approximately 150 ppm, where it was expected to be found. The diol had been prepared previously using the same protocol as above to give a quantitative yield, with the DMT group added prior to the phosphoramidite, again using the same procedure as for compound 1.

### 3.2.3 Compound 3 (n1 on experimental data)

A DMT moiety was attached to one of the alcohol groups on 1,1'-ferrocenedimethanol, as described previously, for the compounds mentioned above. The phosphoramidite was then attached to the alcohol on the other Cp ring, as with the other compounds and characterised by $^{31}$P NMR, which showed that the desired product was present. A quantitative yield was obtained.

## *3.3 Synthesis of Metallocene Nucleobase Precursor.*

Having successfully synthesised the simpler phosphoramidite compounds, a synthetic route to the target ferrocene nucleobase could be developed. One idea was to protect an alcohol group and lithiate the Cp rings, giving the tetra-substituted ferrocene, to which a DNA base could be added, and the phosphoramidite chemistry (demonstrated previously) attempted. This route did not require synthesising a chiral amine precursor prior to lithiation. The second idea was to take an amine pre-cursor (both chiral and achiral) and derivatise it (as seen in **Figure 19**) to give the target. The chiral amine shown below, compound 4, can be synthesised from 1,1'-diacetylferrocene[43]. The Cp rings could then be further substituted to give the ferrocene nucleotide, TfcT.

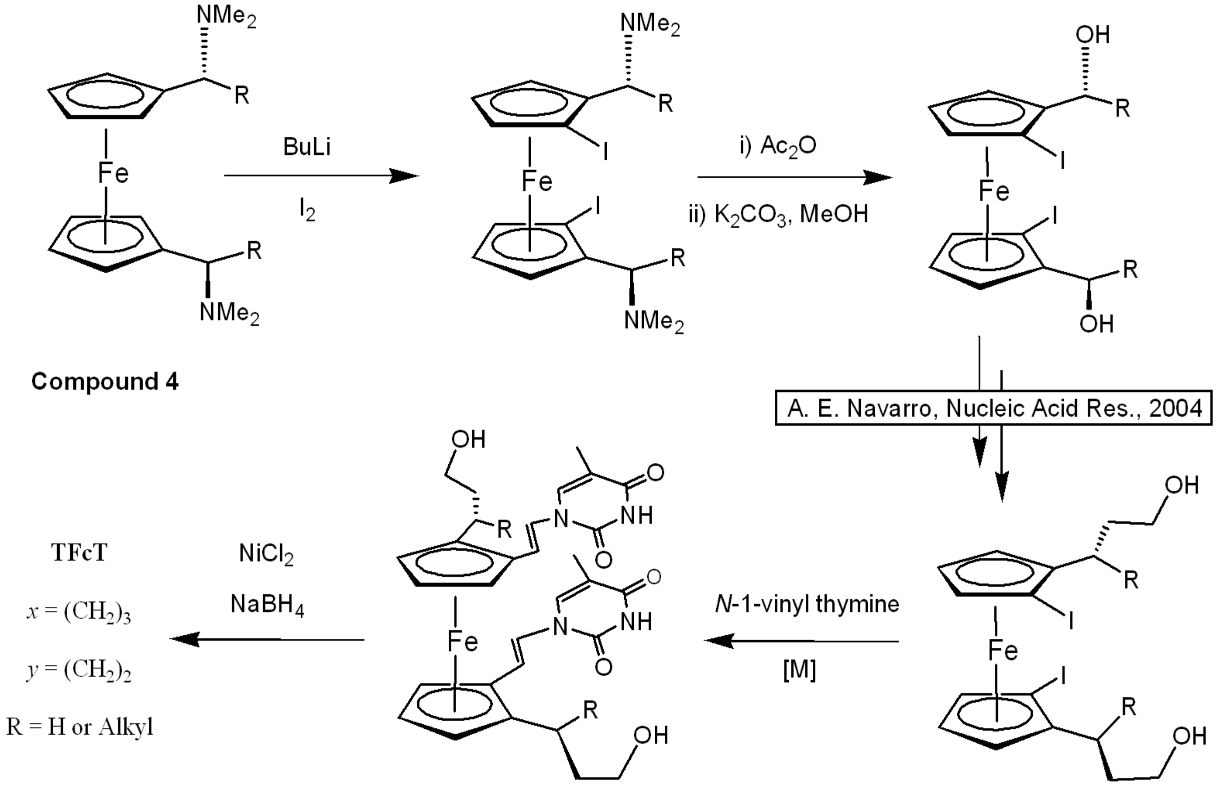

**Figure 19.** Reaction scheme showing the planned steps for synthesis of the ferrocene nucleotide, from the chiral amine (compound 4), to give the x = 3, y = 2 compound, modelled and described in section 1.8. The diol may also be obtained by lithiation of a protected alcohol and subsequent iodination.

The first idea, following on from the methods of synthesising the amides was to lithiate a ferrocene which already had a group attached to each ring (**Figure 20**).

As 1,1'-ferrocenedimethanol was available, it was decided to protect the alcohol and attempt to lithiate the ferrocene, which could then be functionalised with the relevant DNA bases. This would put the required four functional groups on the ferrocene, and the alcohol would need little further functionalising, once deprotected, except to add the DMT and phosphoramidite moieties. The protecting group tertbutyldimethylsilylchloride (TBDMS) was used, as it has the generally desirable properties of being easily introduced, stable to a variety of reactions and can be removed under conditions which do not attack other functional groups[56]. These properties would come in useful when further functionalising the Cp ring and for removal, once the required chemistry was completed.

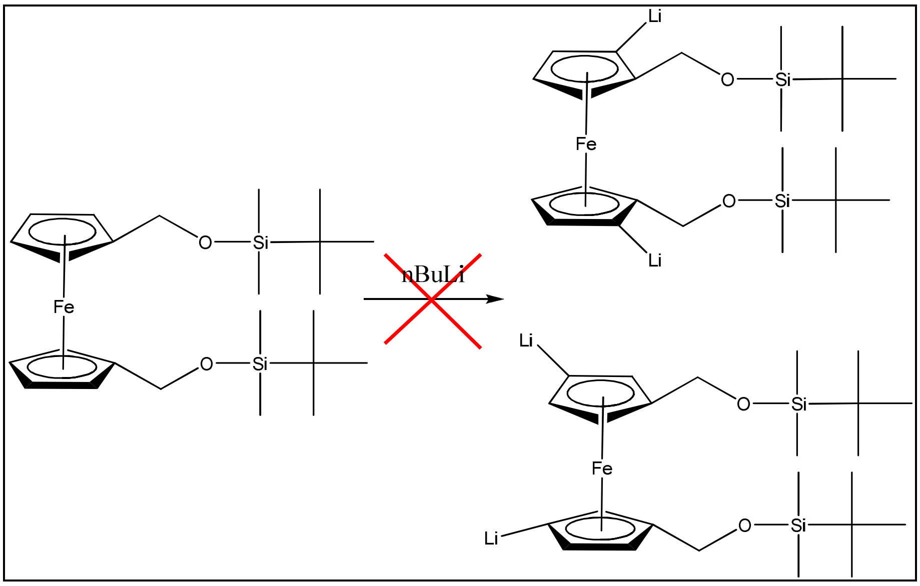

**Figure 20**. Schematic for the lithiation the protected ferrocene, with possible expected outcomes. Lithiation did not take place in practice.

The ease of introduction and removal of TBDMS is influenced by steric factors.[56] As the preferred steric arrangement of the final molecule was not at this time established, it was thought that the functional groups being further apart might be more desirable, giving more flexibility regarding the orientation of the molecule in the DNA strand. This was thought to be a possibility with such a bulky protecting group. It was also desirable to take into account what would be compatible with the addition and cleavage of the DMT group in the oligonucleotide strand synthesis.

The 1,1'-dimethanol alcohol was relatively easily protected using the preparations detailed by Corey.[45] Unfortunately, subsequent lithiation of this molecule proved unsuccessful with n-butyllithium. Ferrocene has long been known to undergo directed diastereoselective ortho-metallation with butyllithium. This reaction was utilised in the synthesis of the phosphoramidite precursors described previously, and was considered appropriate for addition of further functional group to the ferrocene, albeit in a non-directed manner. It is also possible, using lithiation to make a 1,3-substituted ferrocene. This has been shown by Steurer et al..[57]

Three slightly modified protocols were used to lithiate the protected ferrocene. The standard method, as used for undergraduate experiments was the first to be tried, but there was no precipitate formed, indicating that there was no reaction whatsoever.

The second method used was relatively similar, but with a shorter reaction time before addition of chlorodiphenylphosphine with refluxing. However, $^{1}H$ and $^{31}P$ NMR and mass spectrometry showed no evidence of product

The third method was again very similar, but TMEDA was not used. It was postulated that the TBDMS ligand, being bulky, might hinder the formation of the TMEDA adduct required for the previous methods. All other reagents and conditions were kept constant, however, this also gave no result.

Directed lithiation was not demanded of these reactions. The purpose of the reaction was purely to see whether tri- or tetra- substituted compounds could be made from the protected alcohol, but there was no evidence of any metallation at all. When the computer modelling suggested that a product with substitutions in the 1 position (relative to the protected alcohol) would be the optimal orientation for the ferrocene nucleobase, with minimal disruption to the duplex, directed lithiation became more of a concern. As the TBDMS ether is non-directing and bulky (thereby not assuring lithiation in the 1 position) and not particularly effective, it was decided that a different approach might prove more suitable to reach the desired product, as outlined below.

If this route had been continued, the next step would have been to try a stronger lithiating agent, such as tert-BuLi or sec-BuLi. However, this would again not be directed by the TBDMS ether, which was now a requirement to avoid lengthy separations of isomers.

It has been shown that compounds with hydroxyl groups next to the ferrocenyl moiety can be substituted and converted to the corresponding amines, with full retention of configuration. Ugi demonstrated this in 1970[58] when the so-called Ugi's amine was first synthesised. Ugi's amine is shown in **Figure 21.** Once this was successfully made, it was then extended to 1,1'-disubstituted systems, as shown by Schwink and Knochel.[49] This gave a product with two chiral amines attached to the ferrocene. The substitution of an alcohol for an amine, allows ortho-directed lithiation to take place, as the amine is a good directing group for the reaction.

This reaction has been carried out before, as seen in Schwink and Knochel's work, and so established protocols are available.

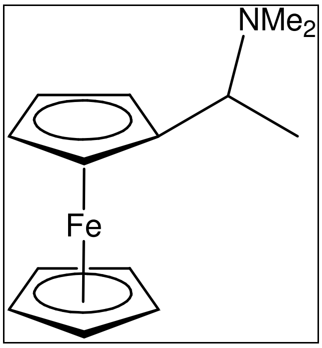

**Figure 21.** Ugi's amine.[58]

It is from this chiral bis-amine compound that the finished product can be synthesised (**Figure 19**). Once lithiated and iodinated, a coupling reaction from the literature could be performed to attach a base to each of the Cp rings.[34] The bis-amine functionality would then be converted a bis-alcohol, with literature procedures followed, to give the TFcT target. This compound could then be functionalised with DMT and phosphoramidite groups, as described earlier.

This synthesis was chosen to follow on from the protected 1,1'-ferrocenedimethanol trials, as it gives the chiral amine product, which can then be separated for insertion into DNA. Ideally, the amine would be enantiomerically pure, so that the lithiation would be stereoselective, resulting in the formation of only one bis-iodo stereoisomer, instead of three.

The separation of the stereoisomers means that there is only one product, which can be incorporated into the DNA. As DNA itself is chiral (having a right-handed helix), a particular isomer may need to be incorporated to preserve the integrity of the helical twist and maintain the stability of the duplex. A mixture of isomers would not give the best yield and may result in problems with synthesis and purification. Using an enantiomerically pure compound would provide some insight into this issue.

The synthesis of the amine involved acylation of ferrocene and then reduction of the resulting ketone group to an alcohol, giving a chiral molecule. An acetate was then attached to the alcohol and dimethylamine used to substitute the O-Acetyl group (**Figure 22**).

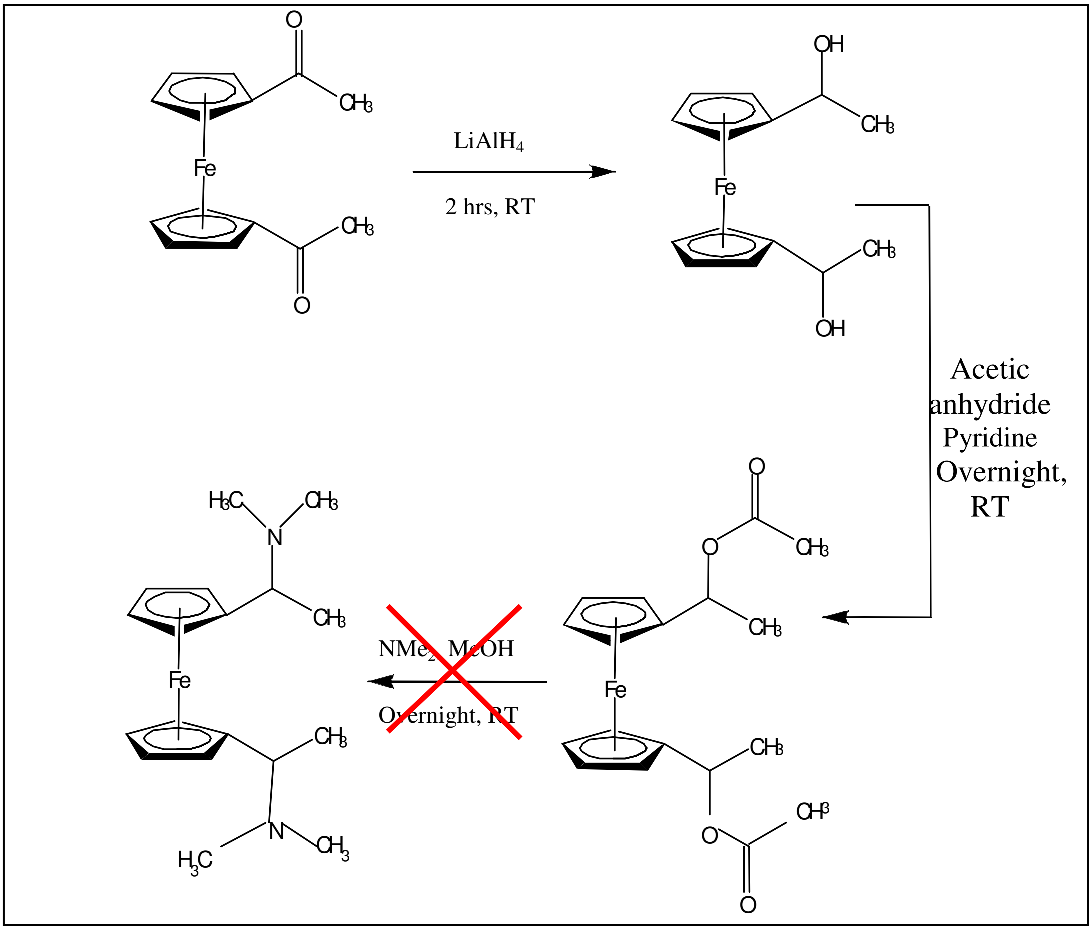

**Figure 22**. Schematic of the reaction steps attempted to get to the $NMe_2$ functionalised ferrocene for use in preparation of the ferrocene nucleobase.

The initial step in this synthesis is the Friedel-Crafts acylation of ferrocene, to yield 1,1'-diacetylferrocene. This was done using a method set out by M. D. Rausch[48] in 1960. This method however, yielded mostly monoacetylferrocene. An alternative method was found,[59] and this gave the diacetyl ferrocene in high purity.

The next step was to reduce the ketone to the chiral alcohol, and this was done using lithium aluminium hydride as the reducing agent, as shown by F. S. Arimoto and A. C. Haven[50]. This was relatively successful, although the yields were not very high, and the product had to be purified by column chromatography. An optimised version of this experiment was found[51], which yielded better results and required no purification steps.

The next step in the synthesis was the acetylation of the reduced ketone. Initially, this was done by using acetic acid and a Dean and Stark water trap[52]. However, the analysis proved inconclusive, and the small quantities of compound being used made this an impractical method.

The next attempt was done following Schwink and Knochel[49], although there were some problems with this procedure. The product required isolation by column chromatography, as the purity was low, but this resulted in all of the product being lost. It also proved difficult to remove the pyridine after the reaction, and so a suitable work up was searched for, as extraction into DCM and washing with water was not productive. The best result was obtained from a procedure used in the same paper for a slightly different compound, requiring more forcing conditions, and so a different work up. The reaction was poured into sodium hydrogen carbonate and extracted into ether, before being washed with water and brine. This made it much easier to remove all the pyridine and resulted in successful isolation of the acetate, which could then be used to make the amine without further purification.

The substitution of the alcohol was also carried out following Schwink and Knochel's work.[49] However, the reaction did not yield any product. An alternative solvent system was tried for a slightly different compound, and this was also unsuccessful. It was therefore not tried with this compound.

A silica chromatography column containing diethyl ether and triethylamine was run to attempt to separate the products and ascertain if a small amount of the amine was present. The NMR results showed that there was no amine present.

An alternative route was then proposed using the 1,1'-ferrocenedimethanol, following the same steps as above, from the acetylation of the alcohol (**Figure 23**).

Using the 1,1'-ferricenedimethanol approach, the isomers would be separated after the iodination and synthesis would then continue as with the 1,1'-dimethylaminoethylferrocene molecule.

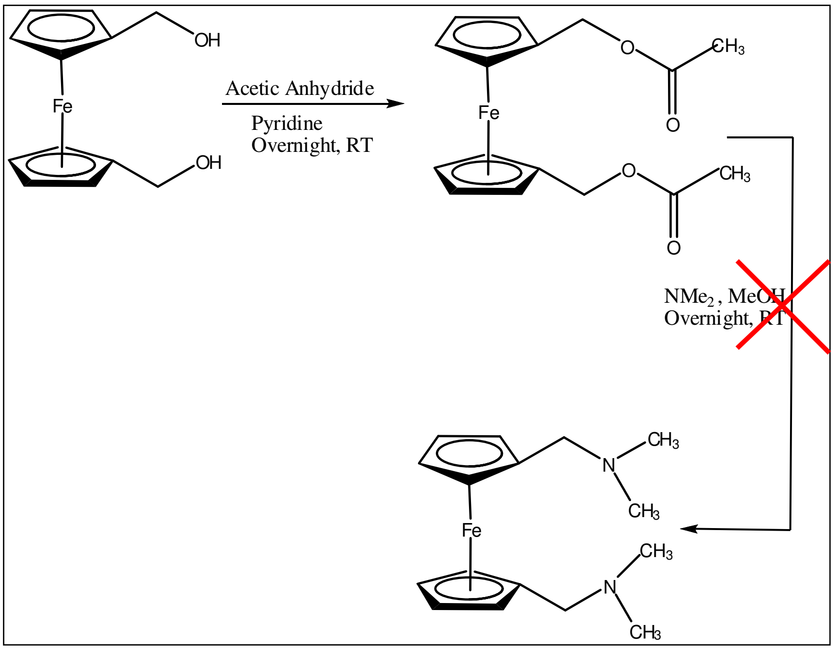

**Figure 23.** Schematic of attempted amine synthesis from 1,1'-ferrocenedimethanol for synthesis of ferrocene nucleobase

This would give a non-chiral bis-amine product which, after the lithiation and iodination steps, would give three different stereoisomers, which would require separation.

The acetylation of this molecule was successful, but the amine substitution was again ineffective. The products recovered were mainly the 1,1'-ferrocenedimethanol, with a small amount of the acetate. As the acetylation reaction for the chiral alcohol was made successful by a small change in the protocol, it was decided to try changing the solvent system to that used for the more reactive substituted ferrocenes in the paper.[49] The methanol system (used on both compounds) can give a methoxy derivative instead of the desired amine, and the mass spectrometry showed masses more similar to this than to the desired amine. The solvent system was changed to THF/water, from methanol/water.

However, this was again unsuccessful, as determined by $^1$H NMR and mass spectrometry. Although study of the reaction by TLC showed that a reaction had occurred, this was the compound returning to the 1,1'-ferrocenedimethanol starting material.

Due to time limitations, no more could be done with the synthesis of either the chiral or non-chiral bis-amine ferrocene nucleobase precursor.

The initial steps of this synthetic route are based on established chemistry, and proved more successful than the previous trial with the protected alcohol. There were some difficulties in obtaining the products, and various methods were attempted to procure the acetoxyethyl group before a successful method was found. A similar problem was encountered with the next step, and time limitations meant that the project could go no further on this point.

If the amine[49] could be obtained however, the three isomers may be separated to give a single chiral product. This stereoisomer could then be lithiated in a stereoselective fashion (as shown in **Figure 19**) and subsequently iodinated.

Having two possible points of enantiomeric separation means that the easiest step at which to separate the isomers can be determined.

### *3.4 Insertion into oligonucleotides and DNA synthesis*

The phosphoramidite compounds described above were inserted into oligonucleotide zzsequences using automated DNA synthesis techniques. The amide with a 3-carbon chain (Compound 2) is referred to as n3 in the sequences and the 1 carbon chain (no amide, Compound 3) is referred to as n1. The first on-column coupling of the ferrocene to the oligonucleotide was not successful, so the coupling was done off column, once the required sequence had been synthesised, using the BTT activator and the phosphoramidite dissolved in anhydrous acetonitrile. The coupling was done using two syringes and the CPG with the desired oligonucleotide sequence and mixing was done for ten minutes and repeated twice.

The sequences chosen for attachment to the ferrocene molecule were very similar, to allow for easy comparison of stability by Tm and UV/ Vis measurements. This made it easier to tell how much more stable a ferrocene on the end of a strand was compared to one placed in the middle of the sequence, or indeed, multiple ferrocenes next to each other in the middle of the sequence.

Various sequences were attempted to be coupled to the phosphoramidites, including ferrocenes inserted into the middle of the sequence, between two guanine bases, and at the end of the sequence, in 3’ and 5’ positions. The sequences used are shown in **Table 1**, with a description of where the ferrocene was inserted.

**Table 1**. Sequences for oligonucleotide synthesis, where the ferrocene was to be attached and how they are referred to in the text. n1 and n3 fer6 were test couplings to the 3' phosphate and are not included in the table

| Sequences attempted | Referred to as | Position of Ferrocene |
|---|---|---|
| 5' FeTGGACTCTCTCAATG | N1 fer1 | 5' end |
| | N3 fer1 | |
| 5' FeCATTGAGAGTGTCCA | N1 fer2 | 5' end |
| | N3 fer2 | |
| 5' CATTGAGFeGTCCA | N1 fer3 | Inserted between two G |
| | N3 fer3 | |
| 5' TGGACTCFeCTCAATG | N1 fer4 | Inserted between two C |
| | N3 fer4 | |
| 5' CATTGAGAGTGTCCAFe | N1 fer5 | 3' end |
| | N3 fer5 | |
| 5' TGGACTFeFeCTCAATG | N1 fer7 | 2 Fc inserted between T and C |
| | N3 fer7 | |
| 5' TGGACTCACTCAATG | Tar MA | Unmodified Strand |
| 5' CATTGAGAGAGTCCA | Tar A1 | Unmodified Strand |
| 5' TGGACTCTCTCAATG | Tar MT | Unmodified Strand |

Examples of oligonucleotides as strands are shown in **Figure 24** and **Figure 25**.

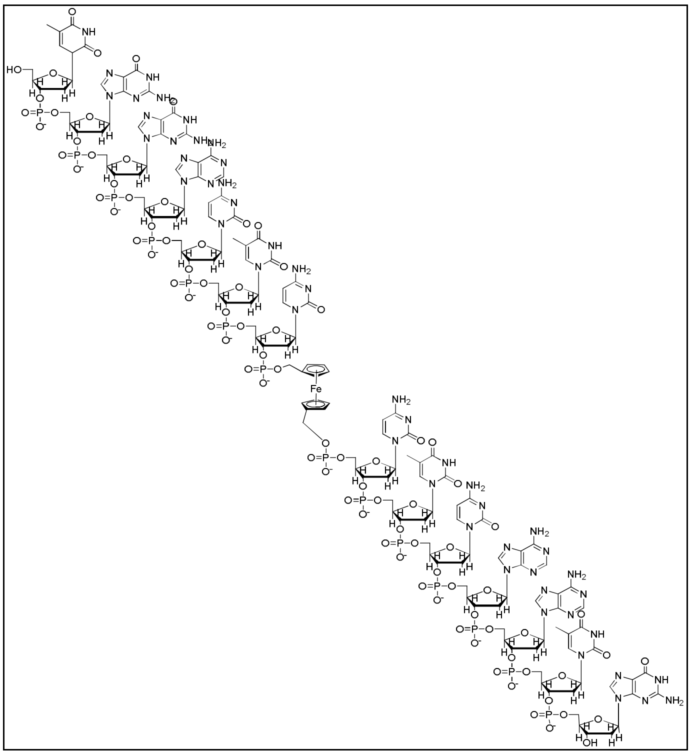

**Figure 24.** Oligonucleotide strand n1 fer3 with the 1,1'- ferrocenedimethanol phosphoramidite in the centre of the sequences.

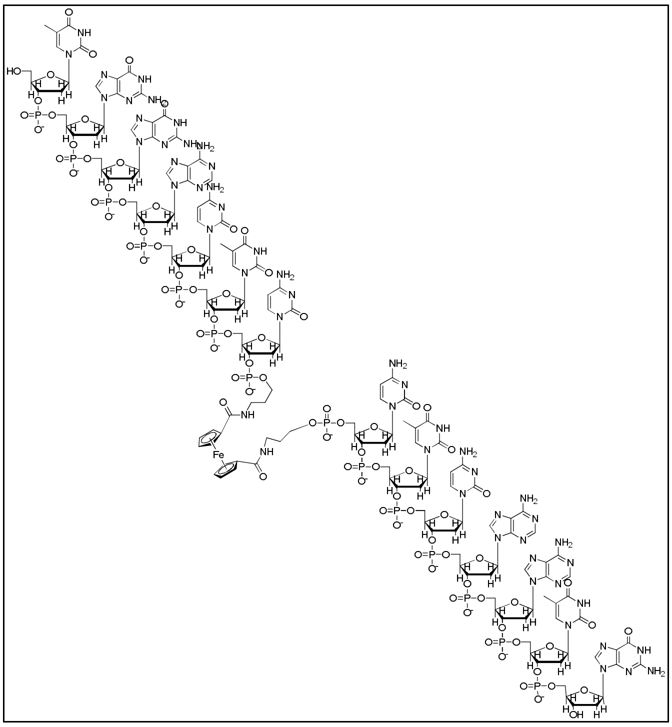

**Figure 25.** Oligonucleotide strand n3 fer3 with the n=3 amide phosphoramidite in the centre of the sequences.

The sequences were purified by reverse phase HPLC. In general, a compound was eluted between 0 and 16 minutes, shown as a broad peak on the traces. This was the DMT off, deprotected, form of unfinished oligonucleotide sequences and by-products, which indicated a failed coupling. The appearance of two main peaks, as seen in some of the traces, shows the broad peak at 12 to 16 and then a sharper, narrower one between 20 and 25 minutes, which was assumed to be the desired oligonucleotide (**Figure 26**). Another reason for multiple peaks is probably that some of the oligonucleotide remained unmodified after the off column coupling, and this was also eluted.

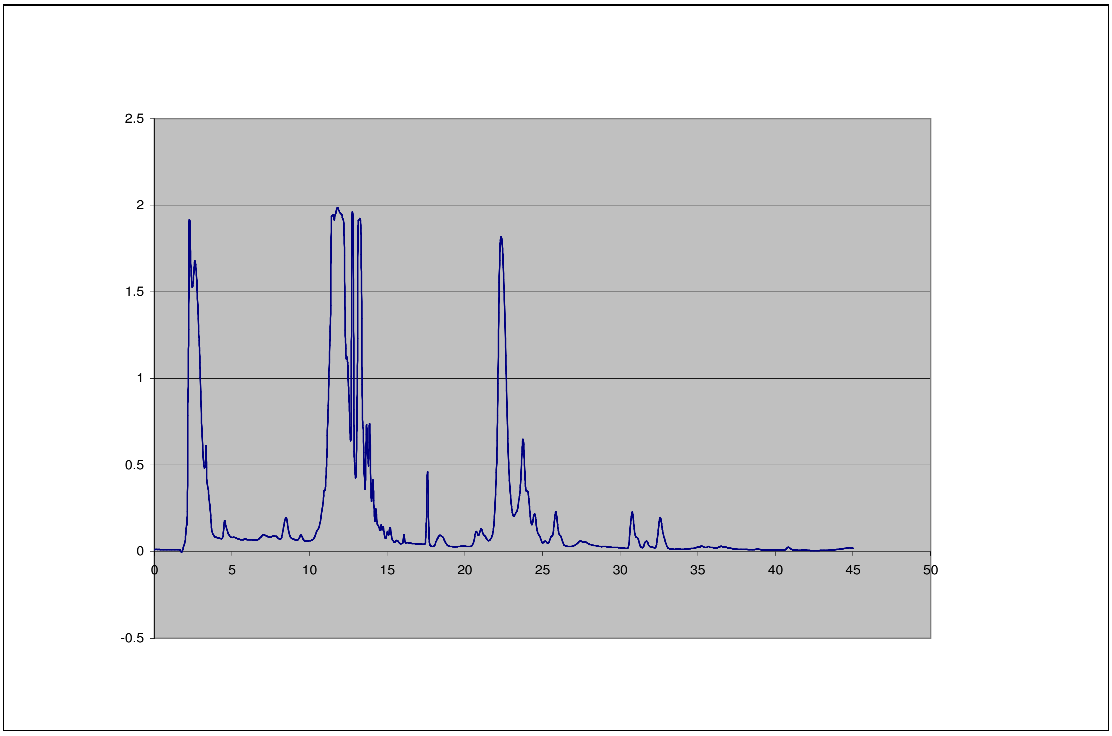

**Figure 26.** HPLC trace showing the main peaks at 15 and 23 minutes.

The main fractions were collected and desalted for the purposes of MALDI mass spectrometry.

Mass Spectrometry is a useful technique, as it can provide structural information about a product. It involves ionisation of the compound being studied, an analyser through which the ionised compound passes and a detector. However, the particular ionisation method needs to be chosen carefully, as it can be difficult to transfer large biological molecules into the gaseous phase without degrading the sample so much as to render it useless. There are various ionisation methods, under two categories: soft and hard ionisation methods. Hard ionisation methods use a lot of energy and tend to result in fragmentation of the molecule, whereas the soft ionisation methods are less destructive, giving fewer fragmentation patterns.

Molecules such as DNA and proteins can be denatured at relatively low vapour pressures and with little variation in pH, making it difficult to get a sufficient amount of the molecule (which must also be an ion) into the vacuum.[18]

Matrix Assisted Laser Desorption Ionisation (MALDI) is an ionisation method better suited to biological molecules, as it is a soft ionisation method (non-destructive). With this technique, the molecule is embedded into a crystal of a matrix (a weak organic acid). The matrix is vaporised by a short laser pulse, and carries some of the molecule embedded into it into the gaseous phase (without excess degradation). Exchange of electrons and/ or protons with the matrix causes ionisation of the target molecule, which can then be analysed[18]. This technique is used in conjunction with a time-of-flight (TOF) analyser, in which heavier molecules take longer to arrive at the detector.

UV quantification was carried out to determine the concentration of the purified modified oligonucleotides. UV quantification of the ferrocene monomers before incorporation into the oligonucleotide sequences was performed, so that the concentration of each strand could be estimated from the absorbance at 260 nm. This allowed comparison before and after incorporation to ascertain if the UV profile was different after coupling to the strand (**Figure 27**). Concentrations can be calculated from the absorbance data by using the Beer-Lambert Law (**Equation 1**). The molar extinction coefficient needed to be calculated to obtain a concentration and the absorbance was measured at 260nm, as this is where DNA absorbs. From this, the percentage yields of the oligonucleotide synthesis could be calculated (**Table 2**).

$$\mathbf{A = \varepsilon c l}$$

**Equation 1**. The Beer-Lambert Law. Where A is absorbance, ε is the molar extinction coefficient, c is concentration and l is path length (usually 1cm)

Unmodified strands were also synthesised and quantified for comparison, to better judge whether the synthesis of the modified oligonucleotide was successful.

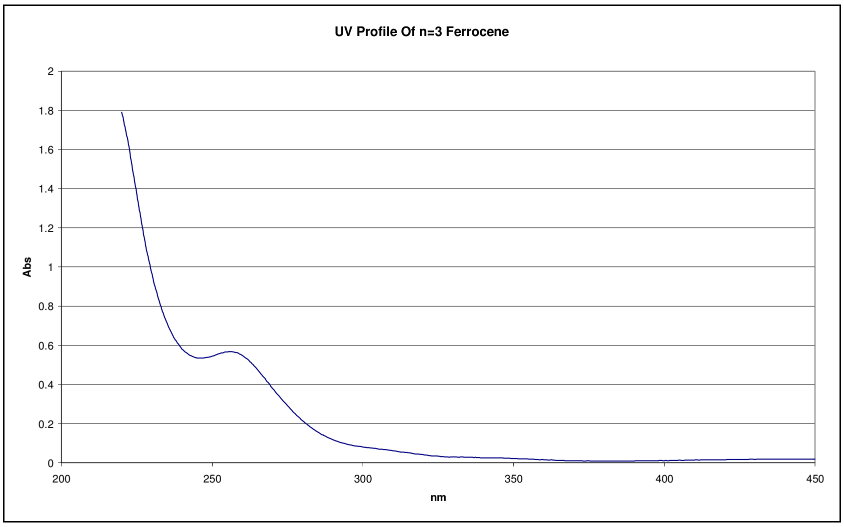

**Figure 27**. The UV profile of the n3 ferrocene monomer (with DMT), ε = 7987.665, before addition of phosphoramidite and insertion into an oligonucleotide strand.

The samples were then made up to a known concentration for $T_m$ studies (melting curves are shown in **Figure 28**), where comparison with the unmodified strands would give an idea of the effect of a ferrocene phosphoramidite on the stability of the strand (see **Appendix 1**).

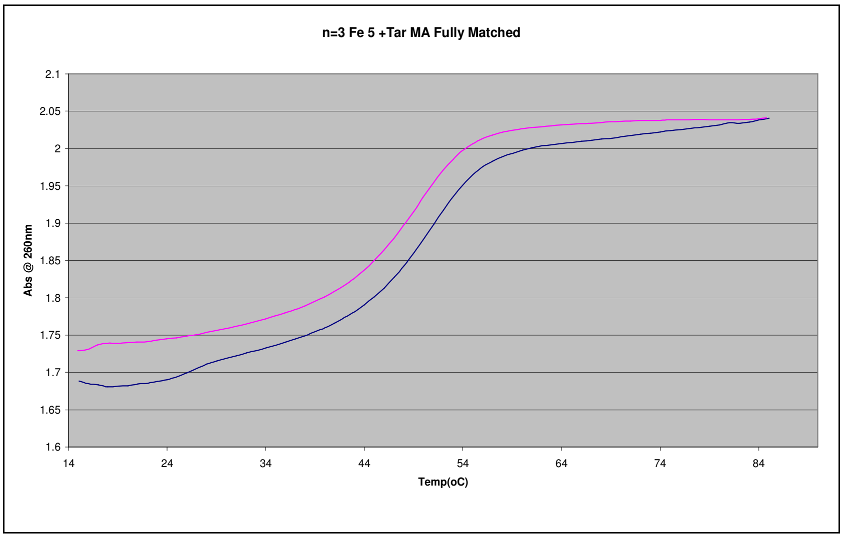

**Figure 28.** Graph showing a melting curve for the n3 fer5 oligonucleotide strand. The two lines show the two melting temperatures measured for the run (ramp 1 and ramp 3).

Of the sequences attempted, n1 fer3, n1 fer5, n3 fer1, n3 fer2, n3 fer4, n3 fer5 and n3 fer7 were deemed most likely to have been successfully modified, from HPLC. The sequences, with the position of the ferrocene in the strand, are listed above (**Table 1**) with UV and $T_m$ data shown below in **Table 2** and **Table 3**. The $T_m$ of unmodified oligonucleotide duplex was calculated to be 52.6 °C. The stability of all the modified strands was only slightly affected by the insertion of the ferrocene phosphoramidites into the sequence, with the most stable being the amides n3 fer4 and n3 fer1 (also with the best yield) annealed with an unmodified strand, although these were more stable than the unmodified DNA. This is unexpected, as there are no π- interaction contributions from the ferrocenes in the strands. However, there may be some hydrogen bonding interactions from the amide.

**Table 2.** Table showing UV data and concentrations determined using it. Samples were dissolved in acetonitrile.

| **Oligonucleotide strand** | **Abs 5 μl (@260nm)** | **Abs 10 μl (@260nm)** | **Average Abs (@260nm)** | **ε / M$^{-1}$cm$^{-1}$** | **concentration/ μmol/ L** | **Yield/ %** |
|---|---|---|---|---|---|---|
| Tar A1 | 96.86 | 165.78 | 131.32 | 153900 | 853.27 | 43 |
| Tar MA | 51.66 | 54.43 | 53.05 | 142300 | 372.78 | 19 |
| Tar MT | 28.93 | 29.82 | 29.37 | 138000 | 212.86 | 11 |
| n=3 Fe 1 | 138.85 | 117.74 | 128.30 | 145988 | 878.81 | 44 |
| n=3 Fe 2 | 0.00 | 65.98 | 32.99 | 157388 | 209.59 | 10 |
| n=3 Fe 4 | 18.40 | 21.70 | 20.05 | 137888 | 145.39 | 7 |
| n=3 Fe5 | 15.86 | 13.17 | 14.52 | 157388 | 92.24 | 5 |
| n=1 Fe1 | 12.70 | 0.00 | 6.35 | 684102.1 | 9.280906 | 0.5 |
| n=1 Fe 3 | 38.86 | 45.40 | 42.13 | N/A | N/A | N/A |

**Table 3.** Table showing the $T_m$ values for the modified oligonucleotides attached to complementary strands. Temperature range is between 15-85 °C

| Strand 1 | Strand 2 | Tm 1/ ℃ | Tm 2/ ℃ | Average Tm / ℃ |
|---|---|---|---|---|
| n3 fer 5 | Tar MA* | 51.1 | N/A | 51.1 |
| n3 fer1 | Tar A1* | 60.1 | 60.1 | 60.1 |
| n3 fer 1 | n3 fer5 | 49.2 | 50.2 | 49.7 |
| n3 fer4 | Tar A1* | 59.3 | 59.3 | 59.3 |
| n3 fer5 | Tar MA8 | 49.9 | N/A | 49.9 |
| *TarA1 and Tar MA are unmodified strands | | | | |

Circular dichroism studies were also carried out. Circular dichroism is a spectroscopic technique, measuring the absorptions of left-handed polarised and right-handed polarise light and comparing the differences in the two (usually very small). Nucleic acids have distinct spectra for certain types of secondary structure, and so structural information can be found using this method[18].

The studies showed some evidence of a chiral signal from the ferrocene, and a definite shift from the unmodified, fully matched strands. The most prominent difference was with an n3 fer4 + TarA1 duplex, which showed a signal from the ferrocene and a definite shift to the left of the spectrum. This indicates that there may be a slight change in the secondary structure of the duplex when a strand is modified with DNA. The spectra are shown in **Figure 29**.

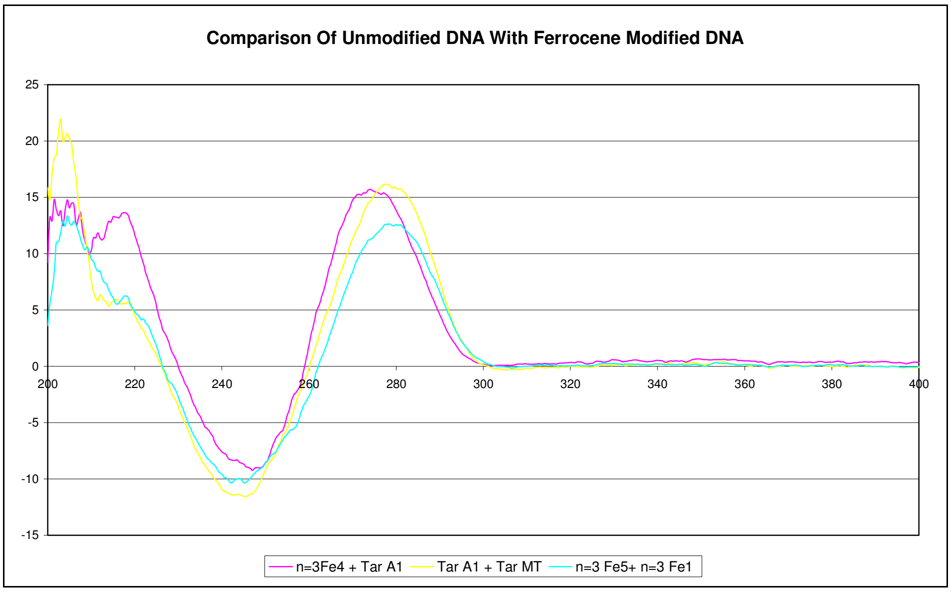

**Figure 29.** Graph showing the CD spectra of unmodified and modified (with one Fe and target and with two modified strands) oligonucleotides. TarA1+TarMT is a fully matched duplex.

Despite appearing to have been successful, the MALDI results for the modified strands were lower than expected, corresponding more closely with the molecular weights of the unmodified strands. This suggests that the incorporation of the ferrocene phosphoramidite monomers was not achieved.

However, the monomers were successfully synthesised, and further attempts to incorporate them could prove more successful, with different chain lengths maybe providing better results.

If the strands were successfully synthesised and used to form a duplex with a complementary strand, also modified, it would be thought that there would be some loss of stability, as the cp rings were not functionalised with DNA bases, or anything that would be able to contribute non-covalent interactions similar to those of the other nucleotides. The measure of stability was attempted, with only one set of strands, due to time constraints, and showed little loss of stability (49.4 °C $T_m$). This would be expected if the strands were actually unmodified. An anthracene-modified strand was also paired with one of the ferrocene-modified oligonucleotides and showed, again, only a small loss in stability, although there were two peaks on the spectrum. This could introduce some interesting possibilities with fluorescent and electrochemical sensing methods, and the use of an anthracene-modified strand provides another type of modification to compare the ferrocene with. It is theoretically possible to put more ferrocene moieties in, as the rest of the strand would be able to form hydrogen bonds with the complementary bases on the other strand as normal, but at a certain point, it would become unstable, as there would be too few non-covalent interactions.

# 4 CONCLUSION

Phosphoramidite monomers for incorporation into an oligonucleotide strand were successfully synthesised and incorporated into a DNA sequence. Automated DNA synthesis techniques were used to synthesise oligonucleotide strands, although there was limited evidence of successful modification of the strands

Future work will build upon this and move closer to the successful synthesis of a ferrocene nucleotide. There are many directions in which this can go, with variations in the bases attached, the number of molecules incorporated into an oligonucleotide sequence, synthesis of an RNA strand and even the metallocene used as the backbone of the molecule.

Although initial attempts to synthesise a ferrocene with four functional groups attached by following a protection chemistry protocol were unsuccessful, they made it possible to move onto the next route. This initial protection approach was relatively simplistic, and so worth trying. However, having ruled out this method, the second approach, using diacetyl ferrocene as a starting material, was devised. This would appear to be more feasible as it allows for a more directed lithiation of the Cp rings.

The reactions completed in the project have been repeated and the conditions altered to optimise them, improving yields and giving a solid base to build on in the future.

Although incorporation of a ferrocene dinucleic acid was not achieved in this project, significant steps have been taken towards completing the aim.

# 5 FUTURE WORK

Initially, to follow on from the work done this year, iodination of the 1,1'-bis(α- N,N-dimethylaminoethyl)ferrocene molecule and progression to the addition of bases can be researched. This will continue the progression towards the final molecule.

In tandem with this, different chain lengths and types of spacers attached to a ferrocene can be researched, to tailor the molecule for incorporation into a sequence. The synthesised modified strands could then also be hybridised to the complementary strand to suggest the feasibility of a duplex forming with ferrocene modified oligonucleotide strands. This research will include stability studies and also, possibly, electrochemical studies.

Once the final molecule has been synthesised and the reaction optimised to give better yields and good incorporation, it may be possible to attach a variety of bases to the ferrocene and do a comparison study of the stability and redox properties of the different molecules. This could then be further extended to comparison of the bases in an RNA type oligonucleotide and of the U and T bases in RNA and DNA strands.

# 6 REFERENCES

(1) Kealy, T. J.; Pauson, P. L. *Nature* **1951**, *168*, 1039.

(2) Miller, S. A.; Tebboth, J. A.; Tremaine, J. F. *J. Chem. Soc.* **1952**, *74*, 632-635.

(3) Fischer, E. O.; Pfab, W. Z. *Z. Naturforsch* **1952**, *7b*, 377-379.

(4) Wilkinson, G.; Rosenblum, M.; Whiting, M. C.; Woodward, R. B. *J. Am. Chem. Soc.* **1952**, *74*, 2125.

(5) Dunitz, J. D.; Orgel, L. E. *J. Chem. Phys.* **1955**, *23*, 954.

(6) Eiland, P. F.; Repinsky, R. *J. Am. Chem. Soc.* **1952**, *74*, 4971.

(7) Moffitt, W. *J. Am. Chem. Soc.* **1954**, *76*, 3386-3392.

(8) Woodward, R. B.; Rosenblum, M.; Whiting, M. C. *J. Am. Chem. Soc.* **1952**, *74*, 3458-3459.

(9) Rosenblum, M. *Chemistry of the Iron Group Metallocenes: Ferrocene, Ruthocene, Osmocene*; Interscience, 1965.

(10) Hagihara, N.; Kumada, M.; Okawara, R. *Handbook of Organometallic Compounds*; W. A. Benjamin 1968.

(11) Coates, G. E. *Organometallic compounds*; Methuen and Co., 1960.

(12) Newman, J. D.; Turner, A. P. F. *Biosens. Bioelectron.* **2005**, *20*, 2435-2453.

(13) Wilkinson, G.; Page, J. A. *J. Am. Chem. Soc.* **1952**, *74*, 6149-6150.

(14) *Ferrocenes: Ligands, Materials and Biomaterials*; Stepnicka, P., Ed.; John Wiley and Sons Ltd., 2008.

(15) Dive, D.; Biot, C. *ChemMedChem* **2008**, *3*, 383-391.

(16) Neuse, E. W. *J. Inorg. Organomet. Polym. Mater.* **2005**, *15*, 3-32.

(17) Top, S.; Vessieres, A.; Cabestaing, C.; Laios, I.; Leclercq, G.; Provot, C.; Jaouen, G. *J. Organomet. Chem* **2001**, *637-639*, 500-506.

(18) van Holde, K. E.; Johnson, W. C.; Ho, P. S. *Principles of Physical Biochemistry*; Pearson International Edition ed.; Pearson Education Ltd., 2006.

(19) Drummond, T. G.; Hill, M. G.; Barton, J. K. *Nat. Biotechnol.* **2003**, *21*, 1192-1199.

(20) McKusick, V. A.; John Hopkins University, Baltimore, MD;Online Mendelian Inheritance In Man (OMIM) 22/7/2008;http://www.ncbi.nlm.nih.gov/entrez/dispomim.cgi?id=602421

(21) Gobbi, M. D.; Viprakasit, V.; Hughes, J. R.; Fisher, C.; Buckle, V. J.; Ayyub, H.; Gibbons, R. J.; Vernimmen, D.; Yoshinaga, Y.; Jong, P. d.; Cheng, J.-F.; Rubin, E. M.; Wood, W. G.; Bowden, D.; Higgs, D. R. *Science* **2006**, *312*, 1215.

(22) Nakatani, K. *ChemBioChem* **2004**, *5*,
1623-1633.

(23) F. Patolsky; E. Katz; Willner, I. *Angew. Chem. Int. Ed.* **2002**, *41*, 3398-3402.

(24) Gooding, J. J. *Electroanalysis* **2002**, *14*, 1149-1156.

(25) Kraatz, H.-B. *J. Inorg. Organomet. Polym. Mater.* **2005**, *15*, 83 - 106.

(26) Ihara, T.; Maruo, Y.; Takenaka, S.; Takagi, M. *Nucleic Acids Res.* **1996**, *24*, 4273-4280.

(27) Sato, S.; Fujii, S.; Yamashita, K.; Takagi, M.; Kondo, H.; Takenaka, S. *J. Organomet. Chem.* **2001**, *637-639*, 476-483.

(28) Sato, S.; Takenaka, S. *J. Organomet. Chem.* **2008**, *693*, 1177-1185.

(29) Ikeda, R.; Chiba, J.; Inouye, M. *e-J. Surf. Sci. Nanotech.* **2005**, *3*, 393-398.

(30) Mucic, R. C.; Herrlein, M. K.; Mirkin, C. A.; Letsinger, R. L. *Chem. Commun.* **1996**, 555-557.

(31) Umek, R. M.; Lin, S. W.; Vielmetter, J.; Terbrueggen, R. H.; Irvine, B.; Yu, C. J.; Kayyem, J. F.; Yowanto, H.; Blackburn, G. F.; Farkas, D. H.; Chen, Y.-P. *J. Mol. Diagn.* **2001**, *3*, 74-84.

(32) Clever, G. H.; Kaul, C.; Carell, T. *Angew. Chem. Int. Ed.* **2007**, *46*, 6226-6236.

(33) Tanaka, K.; Shionoya, M. *J. Org. Chem.* **1999**, *64*, 5002-5003.

(34) Navarro, A.-E.; Spinelli, N.; Moustrou, C.; Chaix, C.; Mandrand, B.; Brisset, H. *Nucleic Acids Res.* **2004**, *32*, 5310–5319.

(35) Inouye, M.; Takase, M. *Angew. Chem. Int. Ed.* **2001**, *40*, 1746-1748.

(36) Beilstein, A. E.; Grinstaff, M. W. *J. Organomet. Chem.* **2001**, *637-639*, 398-406.

(37) Yu, C. J.; Wan, Y.; Yowanto, H.; Li, J.; Tao, C.; James, M. D.; Tan, C. L.; Blackburn, G. F.; Meade, T. J. *J. Am. Chem. Soc.* **2001**, *123*, 11155-11161.

(38) Beaucage, S. L.; Caruthers, M. H. *Tetrahedron Letters* **1981**, *22*, 1859-1862.

(39) Integrated DNA Technologies Inc.: 2005.

(40) Lewis, F. D.; Zhang, Y.; Liu, X.; Xu, N.; Letsinger, R. L. *J. Phys. Chem. B* **1999**, *103*, 2570-2578.

(41) Kashida, H.; Asanuma, H.; Komiyamaa, M. *Chem. Commun.* **2006**, 2768-270.

(42) Mizuta, M.; Terado, T.; Seio, K.; Sekine, M. *Nucleic Acids Symp. Ser.* **2004**, *48*, 237-238.

(43) Department of Chemistry, University of Birmingham: 2008.

(44) Laurent, P.; Miyaji, H.; Collinson, S. R.; Prokes, I.; Moody, C. J.; Tucker, J. H. R.; Slawin, A. M. Z. *Organic Letters* **2002**, *4*, 4037-4040.

(45) Corey, E. J.; Venkateswarlu, A. *J. Am. Chem. Soc.* **1972**, *94*, 6190-6191.

(46) Brown, R. A.; Houlton, A.; Roberts, R. M. G.; Silver, J.; Frampton, C. S. *Polyhedron* **1992**, *11*, 2611.

(47) Hayashi, T.; Yamamoto, A.; Hojo, M.; Kishi, K.; Ito, Y.; Nishioka, E.; Miura, H.; Yanagi, K. *J. Organomet. Chem.* **1989** *370*, 129-139.

(48) Rausch, M. D.; Fischer, E. O.; Grubert, H. *J. Am. Chem. Soc.* **1960**, *82*, 76-82.

(49) Schwink, L.; Knochel, P. *Chem. Eur. J.* **1998**, *4*, 950 - 968.

(50) Arimoto, F. S.; Haven Jr., A. C. *J. Am. Chem. Soc.* **1955**, *77*, 6295-6297.

(51) Hamilton, D. E. *J. Chem. Edu.* **1991**, *68* A143-A144

(52) Gokel, G. W.; Ugi, I. K. *J. Chem. Edu.* **1972**, *49*, 294-296.

(53) Brieden, W. In *http://www.freepatentsonline.com/5760264.html* Lonza AG (CH): Uited States, 1998.

(54) Nakayama, M.; Ihara, T.; Nakano, K.; Maeda, M. *Talanta* **2002**, *56*, 857-866.

(55) Duprey, J.-L.; University of Birmingham Chemistry Deparment: 2007.

(56) Greene, T. W.; Wuts, P. G. M. *Protective Groups in Organic Synthesis*; Third Edition ed.; John Wiley and Sons, Inc., 1999.

(57) Steurer, M.; Wang, Y.; Mereiter, K.; Weissensteiner, W. *Organometallics* **2007**, *26*, 3850-3859.

(58) Marquarding, D.; Klusacek, H.; Gokel, G. W.; Hoffman, P.; Ugi, I. K. *J. Am. Chem. Soc.* **1970**, *92*, 5389-5393.

(59) University of Birmingham; Chemistry Department *Undergraduate Procedure*.

# 7 APPENDIX

## *7.1 DNA Synthesis*

### 7.1.1 Preparation of Buffer A

The triethylammonium acetate buffer (Buffer A) used was made up of milliQ water (800 ml), acetonitrile (50 ml) and acetic acid (5.72 ml). This was made up to pH 6.5 and more milliQ water added to make 1 litre of buffer.

### 7.1.2 De-salting of oligonucleotide sequences

Anhydrous acetonitrile (10 ml) was pushed through a Sep-Pak® cartridge, followed by buffer A (20 ml) to prepare the column. The purified oligonucleotide (in buffer A) was then loaded onto the column and allowed to pass through. More buffer A (5 ml) was pushed through, followed by a wash of water (10 ml). The oligonucleotide was eluted with a 60% aqueous methanol solution (1.5 ml). The solvent was removed in a vacuum centrifuge.

### 7.1.3 UV Quantification, CD and $T_m$ Measurements

The concentrations were determined by UV quantification, which was done by making up a buffer solution (880 µl water, 100 µl phosphate buffer and 20 µl sodium chloride (100 µM) in the cuvette).

The samples were made up to a 5 μM concentration in 10 μM phosphate buffer, at pH 7.0 with 100 μM NaCl.

5 μl of the buffer, with 10 μl of the oligonucleotide, was taken into a microlitre syringe, with 5μl then being dispensed into the cuvette. A reading was taken and the value at 260 nm noted. The rest of the solution in the syringe was added to the cuvette and a second measurement taken and the value at 260nm noted

CD samples were made up as the UV samples.