Pd-catalysed synthesis of functionalised adenosine analogues and their characterisation as base-discriminating fluorescent probes

Sara De Ornellas

PhD Thesis

Department of Chemistry

University of York

September 2012

Abstract

A novel class of 7-modified 7-deazaadenosine nucleosides has been synthesised containing diarylacetylene groups at C7. Diarylacetylene boronate esters were synthesised by chemoselective Sonogashira cross-coupling. The boronate esters were then coupled to 7-iodo-7-deazaadenine nucleosides using aqueous Suzuki-Miyaura cross-coupling in good to excellent yields (68%->99%). The modified nucleosides showed very promising UV-vis absorption and fluorescence emission properties in DMSO, with high quantum yields and UV absorbance maxima bathochromically shifted with respect to intrinsic biomolecular absorption bands. In water, the parent compound **20a** had a decreased quantum yield (0.15) and the absorbance maximum occurred at a lower wavelength (302 nm). The protected phosphoramidite of **20a** was synthesised, and the fluorescent nucleoside was incorporated at a central position in a series of 13-base deoxyoligonucleotides, covering a range of neighbouring DNA sequence contexts. The oligonucleotides had absorbance maxima similar to the nucleoside in DMSO (*ca.* 322 nm), and quantum yields comparable to the nucleoside in water (0.024-0.237). Melting curve analysis of the oligonucleotide duplexes showed some destabilisation due to the fluorescent substituent, but less than expected for a completely mismatched base pair. Circular dichroism spectra highlighted the range of ordered, helix-like structures present in the single-strand oligonucleotides, and confirmed that the duplexes formed a B-type DNA helix. An investigation of base-pair mismatch discrimination showed a clear fluorescence quenching effect for $c^7A\cdot C$ mismatch pairs. Other mismatches did not show discriminatory fluorescence quenching or enhancement. Investigation of the pH dependence of the oligonucleotide fluorescence emission showed high sensitivity to changes in pH, with acidic conditions causing significant quenching. The evidence strongly suggests that mismatched base-pair discrimination is due to formation of a protonated wobble pairing between the 7-deazaadenosine and cytidine, as previously proposed by Seela and co-workers. This base-pair mismatch causes an increase in the pK_a of the 7-deazaadenosine heterocycle, so the mismatch acts as a template for protonation. Potential applications for the 7-modified 7-deazaadenosine nucleoside analogues are discussed.

Contents

List of Schemes

List of Figures

List of Equations

List of Tables

Acknowledgements

Firstly and most importantly, I would like to thank my supervisors Ian and Christoph, for their advice, time and encouragement.

I would like to thank the Chemistry and Biology technical staff, in particular Dr K. Heaton for persevering with MALDI MS, Dr M. Stark for TEM measurements, Dr A. Leech for assistance with CD spectroscopy and Dr R. Sturmey for help with HPLC. Thanks to Dr A. G. Firth for some UV-Vis and fluorescence measurements, and L. S. Griffin for T_m measurements on control oligos. I would like to acknowledge Dr A. Beeby and Bob Edkins (Durham) for TCSPC measurements. I would also like to thank the BBSRC for funding.

Many thanks to all the people who made working in the labs, both in C166 and L1, so much fun. In particular, Tom 1 (Team Nucleoside!), Mike, Jon, Amanda, Tom 2, PhysicsBoy, Tom 3.0 and Little Monkey, for putting up with my ranting, for getting excited about science with me and for always being ready for a pint or a cup of coffee.

Thanks to Maureen, for all the tea.

Finally, thanks to all the people who have supported me along the way. Thanks to my parents, who imparted their love of science, and to Michael for tolerance, understanding, confidence and encouragement.

Author's Declaration

This thesis is a presentation of my original research work, wherever contributions of others are involved, every effort is made to indicate this clearly, with due reference to the literature, and acknowledgement of collaborative research and discussions. The work described within this thesis was performed under the supervision of Prof. Ian J. S. Fairlamb and Dr Christoph G. Baumann at The University of York.

Sara De Ornellas

1   Introduction

1.1   Nucleic acids

Since the 1940s it has been thought that DNA (deoxyribonucleic acid) was the carrier of genetic information. However, it was not until the structure of DNA was determined by Watson and Crick in 1953 that its role and mechanism began to be understood.[1] DNA is a macromolecule consisting of polymerised nucleotides. Nucleotides consist of a nitrogen heterocycle (the nucleobase), a deoxyribose sugar moiety and a phosphate, and the structure of DNA is a sequence of nucleobases joined by a sugar-phosphate backbone. Each DNA macromolecule is paired by hydrogen bonding (through the nucleobases) with a complementary strand, and the two strands run in opposite directions around a central axis to form a double helix (Figure 1.1). There are actually two types of nucleic acids found in cells; DNA and RNA. DNA differs from RNA in that the ribose sugar contains one less hydroxyl, and DNA is double stranded, whereas RNA usually consists of a single strand. In organisms DNA is the storage form of the genetic code, while the transcription of DNA into RNA is required for translation of the genetic code into protein.[2,3]
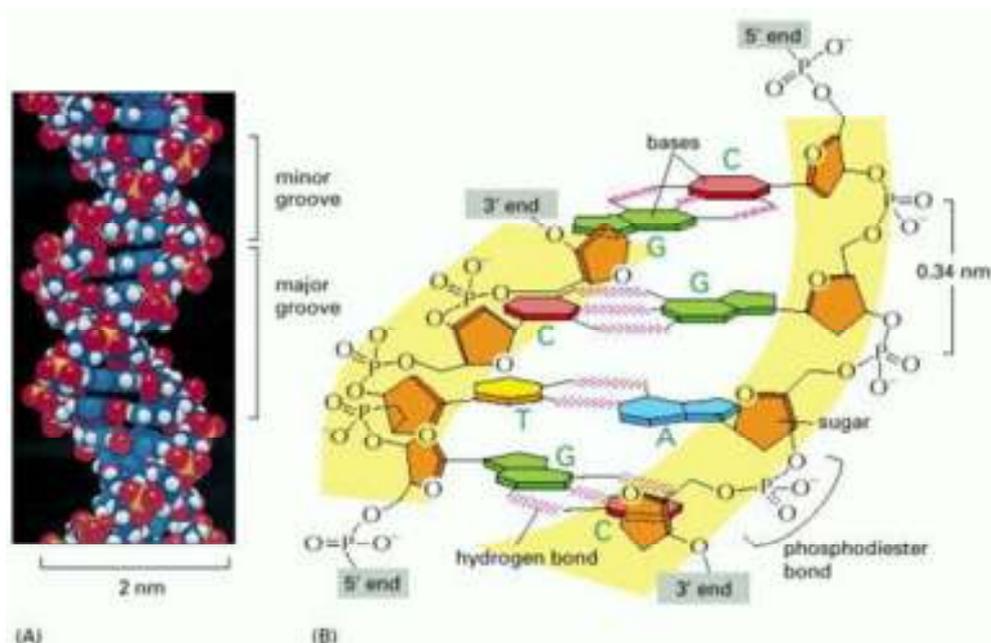
Figure 1.1 Structure of DNA double helix.[3]

1

DNA and RNA each consist of four nucleotides which differ in their nucleobases, three of which are found in both nucleic acids, and one of which is different. The nucleobases can be either purine or pyrimidine heterocycles, and in each case a purine base forms a complementary pairing with a pyrimidine base. DNA generally contains the bases adenine (A), guanine (G), cytosine (C) and thymine (T), whilst in RNA the thymine base is replaced by uracil (U), although other more unusual bases are also found in nature, *e.g.* in transfer RNA. There is specificity in the base pairing (termed Watson-Crick base pairing); adenine pairs with thymine and guanine with cytosine in DNA, with uracil replacing thymine and pairing with adenine in RNA. The base pairs are formed by hydrogen bonding; the A·T/U base-pair has two hydrogen bonds and the G·C base-pair has three hydrogen bonds. The sequence of bases carries the genetic information, and this precise base pairing is crucial to DNA replication, DNA transcription and RNA translation into protein. Nucleobases joined *via* a glycosidic bond to a ribose or deoxyribose sugar are referred to as nucleosides, whilst nucleotides are nucleosides which have been phosphorylated with one or more phosphate esters at the 5'-hydroxyl (Figure 1.2).
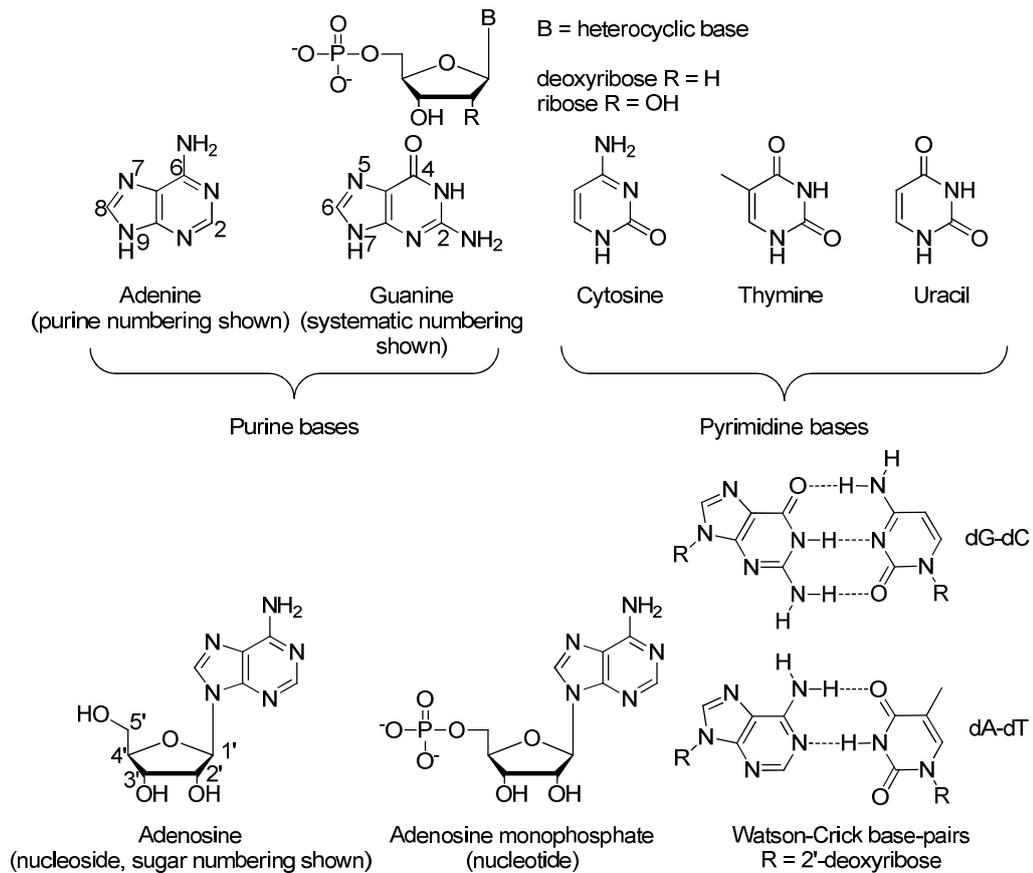
Figure 1.2 Structure of nucleobases, nucleosides and nucleotides.

1.1.1   Nucleic acid secondary structure

Nucleic acids can adopt a variety of helical structures. The most common structure associated with DNA is the B-type double helix, which is shown above. In this structure, the antiparallel strands form a right-handed helix where the stacked base-pairs lie perpendicular to the axis of the helix. The A-type double helix is also a right handed helix, but differs in the geometry of the bases and the conformation of the sugars. In A-type structures, the base-pairs are tilted with respect to the helical axis, and the helix is wider and shorter. A-type helices bind less water than B-type structures, so can be observed crystallographically in dehydrated nucleic acids. RNA double helices do not generally form B-type structures due to steric clashes between the C2′-OH and the closest phosphate group, so double-stranded RNA tends to adopt an A-type structure, where the change in sugar conformation means the C2′-OH projects outwards. A third (and less common) biologically relevant structure is the Z-type helix. Unlike A- and B-type nucleic acids, Z-type structures are left-handed helices. Z-type structures are generally formed by duplexes of alternating guanosines and cytidines. The nucleotide conformations are also very different from A- and B-type structures; rather than all the nucleotides adopting the *anti* conformation, the nucleotides alternate between *anti* and *syn*, with the purines generally adopting the *syn* conformation (Figure 1.1). The key features of these nucleic acid structures are summarised in Table 1.1.
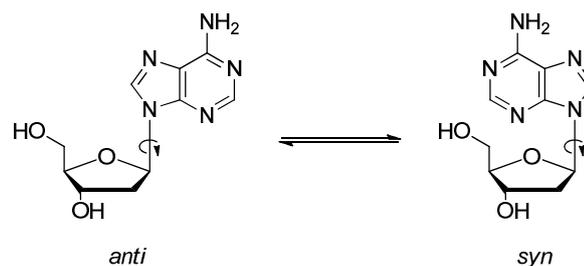
anti   *syn*

Figure 1.3 *Anti*/*syn* nucleoside equilibrium, which defines the position of the purine relative to the sugar ring system.
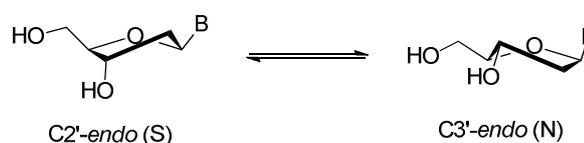
C2'-*endo* (S)   C3'-*endo* (N)

Figure 1.4 C2′-*endo*/C3′-*endo* sugar pucker equilibrium.

Table 1.1 Key structural features of common nucleic acid secondary structures.

	Helix structure		
	A	**B**	**Z**
Helix sense	Right-handed	Right-handed	Left-handed
Glycosyl bond	*anti*	*anti*	Alternating *syn* and *anti*
Rise per base pair /Å	2.3	3.4	3.8
Base-pair tilt with respect to helix axis	19°	1°	9°
Sugar pucker	C3′-*endo* (N)	C2′-*endo* (S)	Alternating C2′-*endo* and C3′-*endo*

The factors which affect the nucleic acid secondary structure formed are diverse. As previously mentioned, hydration of the helix can play an important role, as can the sequence of nucleotides. The environment of the nucleic acid is also key: high salt environments stabilise B-type structures, while addition of less polar solvents such as ethanol or trifluoroethanol can cause transition to A-type structures in which the hydrophobic nucleobases are more exposed. The conformation of the sugar is also very important (Figure 1.4), as well as steric accommodation of any structural modifications.

Other base-pairing modes are also known, which give rise to more varied secondary and tertiary structures.[4] These non-Watson-Crick base-pairs are most commonly observed in RNA structures. Hoogsteen base-pairs are formed along the major groove edge of nucleic acid duplexes, and do not interfere with normal Watson-Crick base-pairs (Figure 1.5). This means that where both Watson-Crick and Hoogsteen base-pairs are present, RNA triplex structures can be formed. In some cases, such as shown in the C·G*A$^+$ structure in Figure 1.5, the nucleobase must be protonated in order to form a Hoogsteen base-pair.

T-A*T   C-G*A$^+$
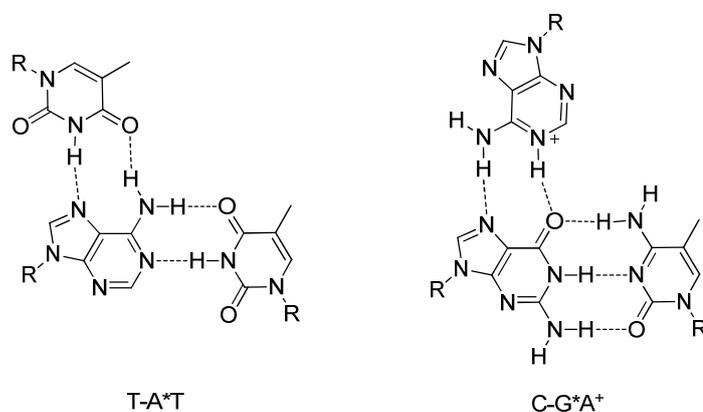
Figure 1.5 Examples of Hoogsteen base-pairing.

1.2 Modified nucleosides and nucleotides

Modified nucleosides have been the subject of much research due to the wide range of biological activity observed in many derivatives.[5] Analogues of the naturally occurring nucleosides, in particular sugar-modified nucleosides, have been developed for medicinal purposes.[6] They have been widely used as antiviral agents, for example as treatment against HIV (Zidovudine and Stavudine) and herpes virus (Acyclovir and Famciclovir) (Figure 1.6).[7] Nucleoside analogues have also been used in the treatment of some cancers.[6]
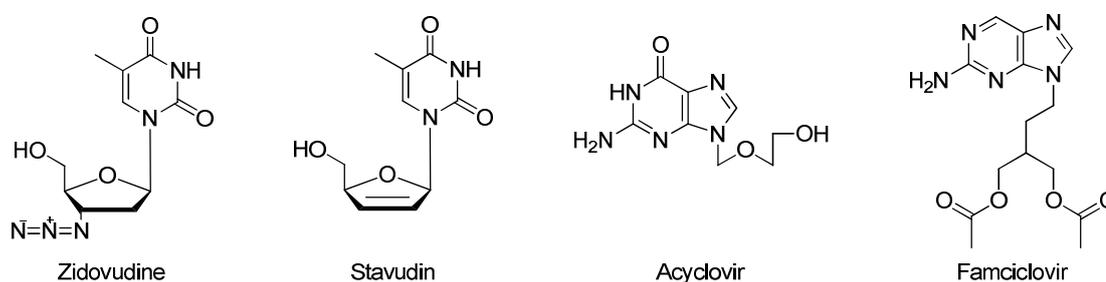
Figure 1.6 Examples of modified nucleosides used as antiviral drugs.

1.2.1 Fluorescence labelling for biochemical applications

Nucleosides and nucleotides can be chemically-modified so that they can be used as probes of biological systems. A variety of spectroscopic labels are commonly used to study biological systems, depending on the specific requirements of the system being investigated and the information required. These include spin-labelling (NMR[8] and EPR[9]), radiolabelling,[10] electrochemical-labelling[11] and fluorescence-labelling.[12,13] Fluorescence-labelling is probably the most widely used, as it is a safe and extremely versatile technique. Fluorescence spectroscopy can be applied to a number of different techniques to determine both structural and dynamic information about a biological system. Both ensemble and single-molecule fluorescence techniques can be used. Ensemble methods look at fluorescence signals averaged over a large population of molecules. One of the great advantages of fluorescence spectroscopy is its high sensitivity, allowing for the detection of a single fluorophore with appropriate excitation and detection methods. Advanced techniques such as FRET (Förster Resonance Energy Transfer) are powerful tools

for investigating 3D structure and molecular dynamics.[14] FRET is where the emitted photon of one fluorophore (the donor) is absorbed by a second fluorophore (the acceptor), giving information about the relative distance of the fluorophores.

Another valuable property of many fluorophores is the dependence of their fluorescence emission on their chemical environment. Environmental sensitivity is usually dependent on the polarity or pH of the location of the fluorophore – for example, a nucleic acid labelled with an environmentally sensitive fluorophore will show different fluorescent emission depending whether or not it is bound to a protein. In this way, fluorophore-labelled biomolecules are powerful tools for biochemical analysis.

1.2.2 Fluorescent nucleotides

It is generally desirable that the functionalised-nucleotides exhibit the same biological properties as the natural compounds.[13] This usually includes base-pairing and duplex stability, but is dependent on the desired application. The fluorescent properties of the modified nucleotide are also important – the introduced fluorophore must absorb light at wavelengths shifted away from the background absorbances of nucleic acids (260 nm) and proteins (280 nm). It is also important that the emission spectra be distinct from the intrinsic fluorescence emission of tryptophan in proteins (300-350 nm). Another consideration when choosing a fluorescent label is the linker which attaches it to the substrate. Fluorophores may be directly attached to nucleotides, or linkers (usually alkyl chains) can be used. The length and nature of the linker can significantly affect how sensitive the probe is to changes in the environment of the nucleotide.

Modifications can be made at almost every position, but can be divided into three main categories: base-modified, sugar-modified and phosphate-modified (Figure 1.7).[15,16]

6
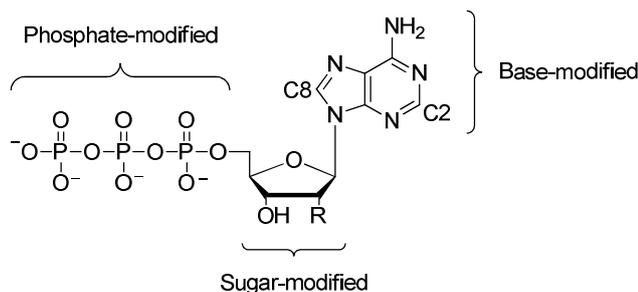
Figure 1.7 Possible positions for structural modification of nucleotides.[*]

Nucleosides and nucleotides can be labelled at either the 2′- or 3′-OH on the sugar, coupled directly or through a linker. These compounds are useful for studying systems where the nucleotide can bind even with large groups on the sugar moiety (for example, where the NTP interacts with proteins through the base, and the sugar points away from the binding site).[17] Modifications at the 3′-OH are not suitable for incorporation into nucleic acids, as the nucleotide is linked to its adjacent nucleotide through the 3′-O position (Figure 1.8). As deoxyribonucleotides do not contain a 2′-hydroxyl, modifications cannot be made directly at this position in DNA systems. It is also important to ensure that modifications at the 2′-H or 2′-OH position do not interfere with the conformation of the ribose 5-membered ring. The sugar conformation determines the secondary structure of the nucleic acid, and substitutions at this position can influence the ratio of conformations both sterically and electronically.[18]
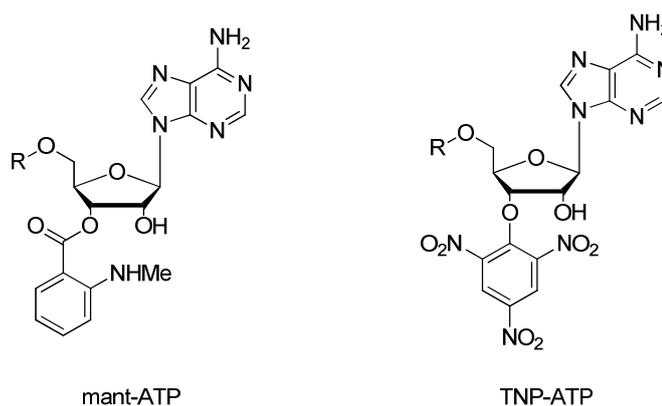
Figure 1.8 Examples of commercially available sugar-modified fluorescent ATP analogues.[16]

---

[*] Purine numbering used throughout main text.

7

Phosphate-modified nucleotides have also been used extensively, particularly for single-molecule DNA sequencing,[19] and can be functionalised in a similar manner (Figure 1.9). However, as γ-phosphate labels are cleaved upon nucleotide incorporation into oligonucleotides, they are less suitable for studying nucleic acids.
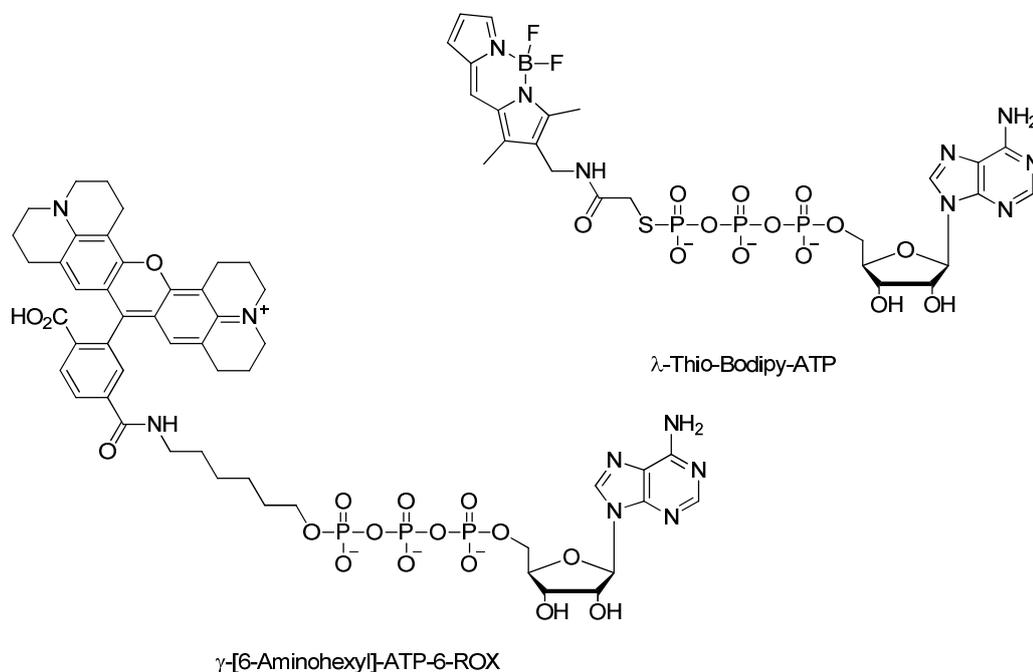
Figure 1.9 Examples of commercially available phosphate-modified fluorescent ATP analogues.[16]

Many base-modified fluorescent nucleotides are also known, and many are commercially available. Generally, the disadvantage of this strategy is the possible interference with Watson-Crick base-pairing. For adenine nucleotides, there are three main positions for attachment of a fluorophore. Labels can be introduced on the C2 position of the nucleobase, however this will interfere with the Watson-Crick edge. There are many applications where this is not problematic; however these modified nucleotides are not suitable for studying DNA duplexes or many RNA systems with secondary structures based on Watson-Crick base-pairing. Labels can also been introduced by substitution at the exocyclic amine. Even labels with long linker groups at this position can influence the Watson-Crick base-pair by altering the electronic properties of the amine proton. Other modifications at this position can block the Watson-Crick base-pairing face altogether, or reduce base-pairing specificity (Figure 1.10).
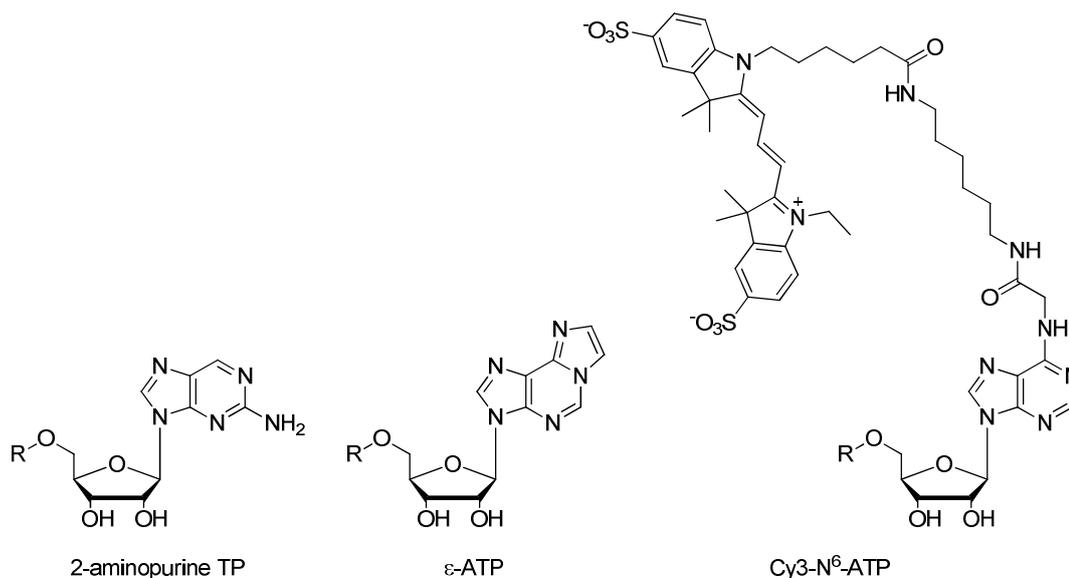
Figure 1.10 Examples of commercially available base-modified fluorescent ATP analogues.[16]

Therefore, an advantageous position for labelling purine nucleotides is at the 8-position, which is removed from the Watson-Crick base-pairing edge; many 8-modified purine nucleosides are also known. However, there is also a major disadvantage to modification at this position. Large groups at C8 (for example phenyl or phenylethynyl, see Section 1.4) cause a switch from the sugar being predominantly in the *anti*-position, to almost exclusively in the *syn*-conformation.[20] This switch is accompanied by a change in the ratio of 2′-*endo* (S) to 3′-*endo* (N), favouring the 2′-*endo* conformation (Figure 1.11).
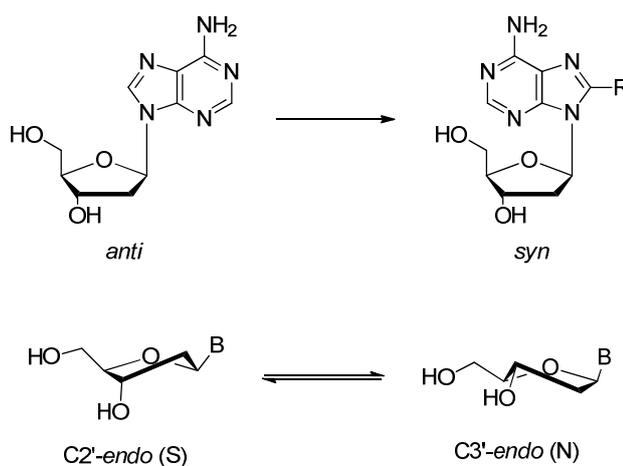
Figure 1.11 Conformational changes which are induced by substitution at C8 of adenosine.

9

1.2.3 7-Deazapurine nucleosides

The conformational changes observed with 8-modified purines led to the development of labelling at C7 of 7-deazapurines. 7-Modified-7-deazapurine nucleotides have been shown to be enzymatically-incorporated into nucleic acids, and show high duplex stability. The 7-substituent is located in the major groove of DNA and therefore does not interfere with stability or base-pairing. Unlike the 8-modified purines, 7-modified-7-deazapurines preferentially adopt the *anti*-conformation.

Naturally occurring 7-deazapurine nucleosides such as tubercidin,[21] sangivamycin[22] and toyocamycin[23] have been of interest for many years due to their biological activity and potential pharmaceutical applications (Figure 1.12). Tubercidin (7-deazaadenosine) was first isolated from *Streptomyces tubercidicus* in 1957, and the structure confirmed as 4-amino-7-D-ribofuranosyl-7*H*-pyrrolo[2,3-*d*]pyrimidine in 1963.[24] [14]C labelling studies suggest that tubercidin is biosynthetically derived from adenosine.[25] C2 of adenosine is conserved in tubercidin, but C8 of adenosine does not become C8 in tubercidin. Instead, it appears that the carbon atoms which make up the pyrrole ring of tubercidin originate from C1 and C2 of ribose (Scheme 1.1).
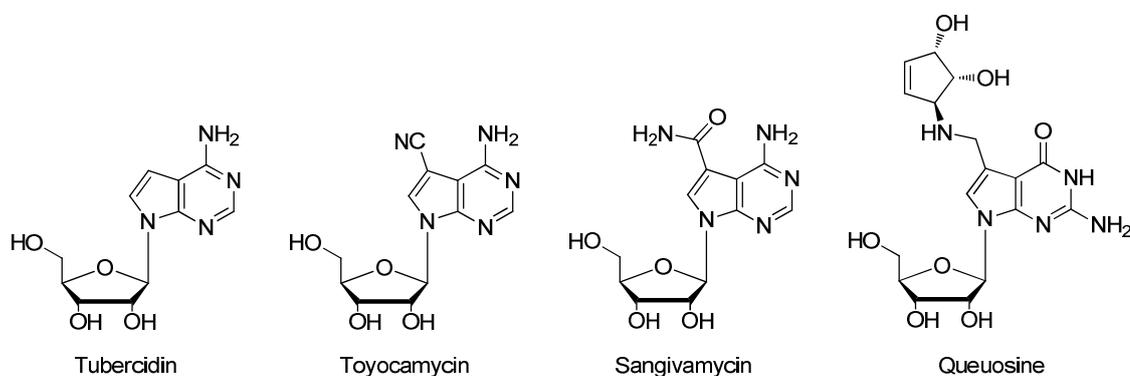
Figure 1.12 Examples of naturally occurring 7-deazapurine nucleosides.

One of the most important naturally-occurring 7-deazapurine nucleosides is the 7-deazaguanosine-derived queuosine (Q) which is found in nearly all prokaryotes and eukaryotes.[26] It replaces the first anticodon position in the tRNAs of histidine, asparagine, aspartic acid and tyrosine; it is thought to enhance translation accuracy, amongst other purposes.
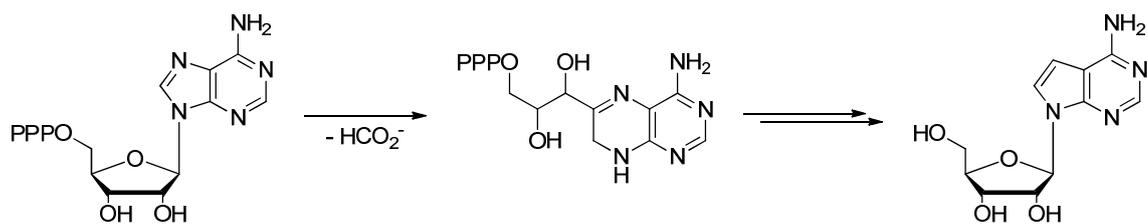
Scheme 1.1 Proposed key step in the biosynthesis of tubercidin.

7-Deazapurine nucleosides can be easily functionalised at the 7-position, and because of their high duplex stability and similar properties to the natural purines, they have become popular choices for the synthesis of labelled purine nucleosides. However, there are surprisingly few examples of usefully fluorescent 7-modified-7-deazapurine nucleotides, i.e. those with appropriate absorbance and emission wavelengths and high quantum yields. The Seela group has published a range of fluorescent 7-deazapurine nucleotides with various organic groups connected by alkynyl linkers.[27] However the quantum yields tend to be very low and the absorbance maxima not sufficiently shifted away from those of biomolecules. The examples shown in Figure 1.13 show the analogue with the highest quantum yield (recorded in water), and the analogue with the highest absorbance maximum. Generally, these analogues have fluorescence emission too weak for most applications in biological systems.

$\lambda_{max} = 280$ nm
$\lambda_{em} = 405$ nm
$\Phi = 31\%$

$\lambda_{max} = 286$ nm
$\lambda_{em} = 412$ nm
$\Phi = 2\%$
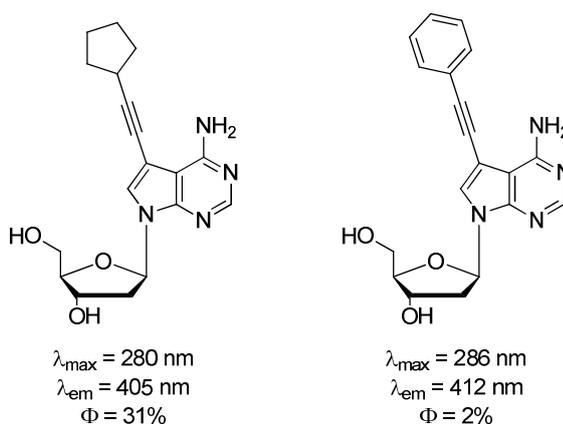
Figure 1.13 Examples of fluorescent 7-modified-7-deazapurine nucleosides reported by the Seela group.

Kaufmann *et al.* published an alternative fluorescently-labelled 7-deazaadenosine, containing a stilbene fluorophore with a PEG (polyethylene glycol) linker (Figure 1.14).[28] They attached an aminopropynyl group to the 7-deazadenosine, to which the fluorophore with PEG tether could be covalently attached using amide chemistry, to give the fluorescent nucleoside which was then phosphorylated. The 7-deazaadenosine triphosphate could be incorporated into DNA by a variety of DNA polymerase enzymes, although the rate of incorporation was quite low. A study by Famulok and co-workers showed that similar modifications to these are better tolerated by family B polymerase enzymes than family A polymerases.[29]
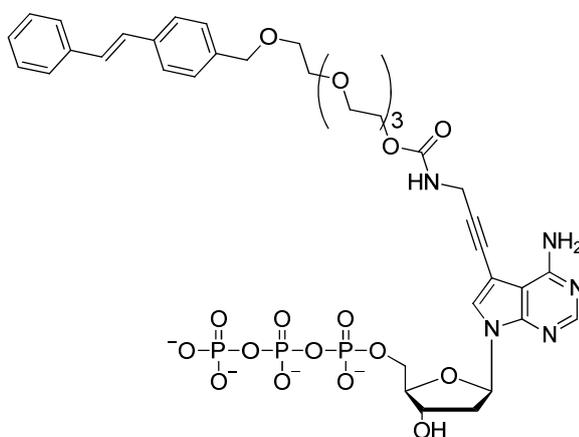
Figure 1.14 7-Deazaadenosine triphosphate with PEG-tethered stilbene fluorophore.

Another example of 7-deazapurine nucleotides containing various fluorophores was reported by Ju *et al.* in 2006.[30] They used alkyne linkers, once again, but this time the fluorophore was connected by a cleavable linker (Figure 1.15). The added flexibility gained by using a long linker allows for the addition of a large fluorophore, giving extremely fluorescent nucleotides (in this case for DNA sequencing applications). The fluorophores (and 3'-O-allyl protecting groups) are cleavable using a Pd catalyst. These fluorescent nucleotides have been shown to be tolerated by DNA polymerase enzymes, and have applications in DNA sequencing using PCR. The disadvantage of attaching a fluorophore through a long linker is that the reporter part of the molecule is distant from the biologically active centre, meaning it will most likely not be sensitive to the microenvironment of the nucleobase. It is also possible for planar fluorophores attached to nucleotides through long linkers to intercalate DNA helices, which is sometimes desired, but may cause fluorescence quenching.[31]
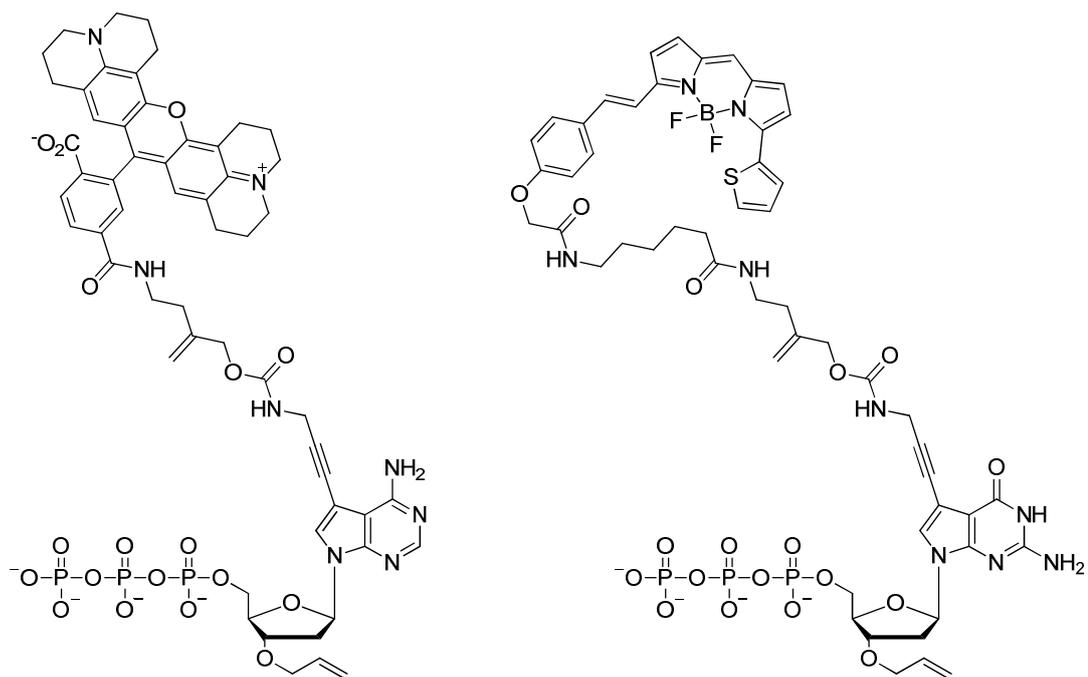
Figure 1.15 Fluorescently-labelled 7-deazapurine nucleotides for DNA sequencing.

Other types of labels for 7-deazapurines have been reported. The Hocek group has reported a variety of nucleosides and nucleotides with electrochemical labels, including 7-deazapurines. The electrochemical labels include Ru-bipy complexes,[32] ferrocene substituents[33] and *m*-nitro- and *m*-aminophenyl labels (Figure 1.16).[34] The compounds utilise both alkynyl and phenylene linkers.
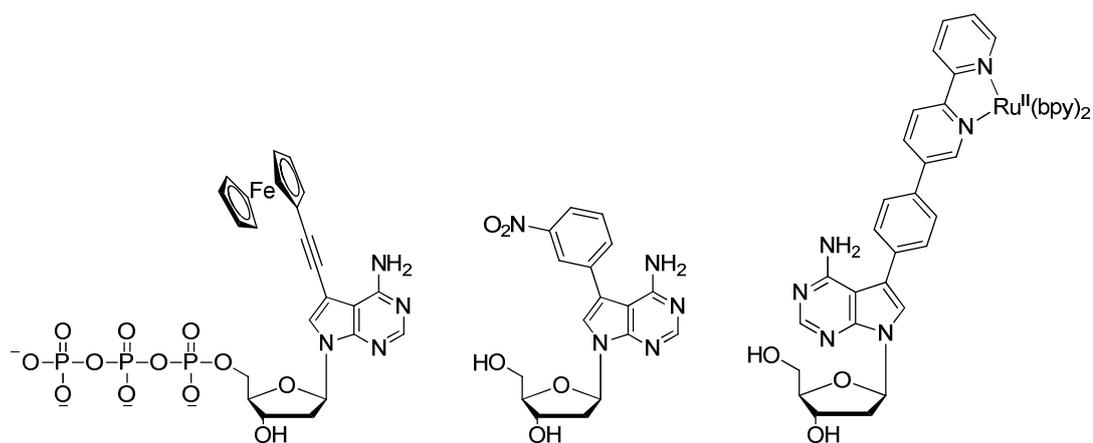
Figure 1.16 Representative electrochemically-labelled 7-deazaadenosine nucleosides and nucleotides reported by the Hocek group.

13

1.3   Fluorescent nucleotides for base mismatch discrimination

1.3.1   Single nucleotide polymorphism

Single nucleotide polymorphisms (SNPs) are site-specific mutations within a gene resulting from replacement of a single nucleotide with a different base. Polymorphisms can be between individuals of a species, or between paired chromosomes of an individual. SNPs are more often found within non-coding sequences of genome, and even within coding regions do not necessarily result in the coding of an incorrect amino acid (due to the degeneracy of the genetic code). Despite this, SNPs (or combinations of SNPs) have been implicated in many human diseases, such as osteoporosis,[35] cardiovascular disease,[36] and primary (idiopathic) hypertension,[37] amongst others. SNPs can also indicate an individual's response to drugs, and therefore may hold the key to future developments in personalised medicine.[38]

1.3.2   Methods for genotyping SNPs

There are numerous methods available for the identification of known SNPs. They can be divided into four main categories: detection of hybridisation, primer extension, ligation and invasive cleavage.[39] Assays can either be carried out in solution (homogeneous) or on a solid support. Invasive cleavage involves the application of the enzyme Flap endonuclease (FEN) which is a very sensitive enzyme for structure-specific cleavage.[40] The Invader® assay involves two oligonucleotide probes which complement different parts of the target sequence – one labelled with a fluorescence quencher on one side of the cleavage site and a fluorophore on the opposite side, and an Invader oligonucleotide. The two oligonucleotide sequences overlap at the single-nucleotide polymorphism site. If the target sequence is fully complementary with the probe strand, cleavage can occur and the fluorophore-containing oligonucleotide is released, giving fluorescence emission (Figure 1.17).
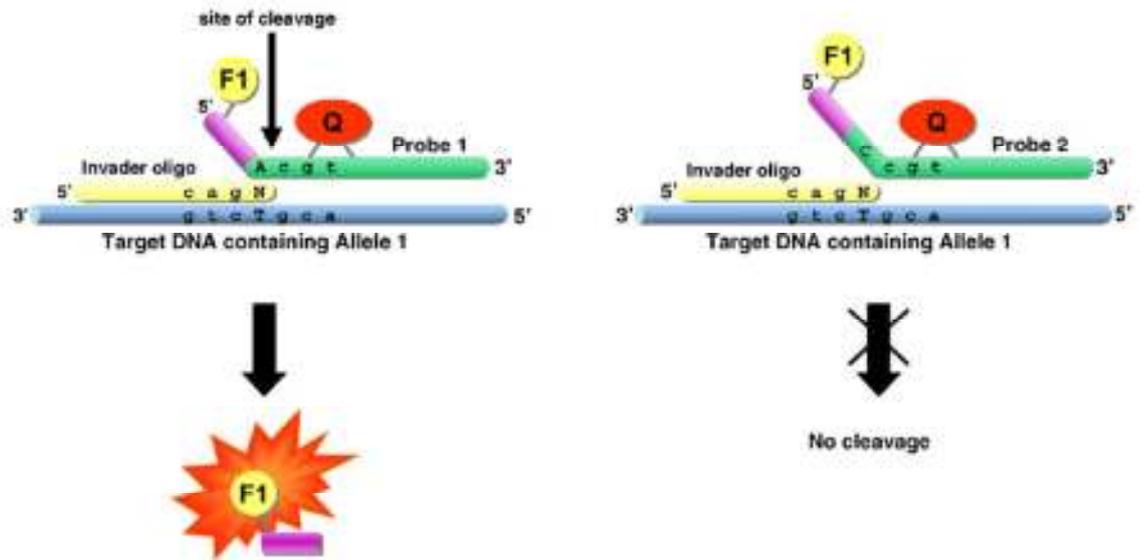
Figure 1.17 Representation of the Invader® assay for SNP genotyping.[40]

The oligonucleotide ligation assay (OLA) also uses two oligonucleotide probes which complement the target sequence.[41] In this case there is no overlap – the two sequences anneal to the target strand and are joined by DNA ligase. Two capture oligonucleotides are used, labelled with different fluorophores, and with different nucleotides at the 3' end, representing the two genotypes. The fluorophore which is successfully ligated represents the allele which is present, and so can be used to identify the genotype.

Primer extension methods involve the use of polymerase enzymes which can extend a primer sequence to incorporate a nucleotide complementary to the SNP.[39] The synthesised sequence can then be detected using a variety of techniques such as mass spectrometry or melting curve analysis. Alternatively, the primer sequence can contain the SNP. In this case, a PCR reaction is used where the sequence is only amplified if the primer is fully complementary to the target sequence.

Hybridisation methods directly detect whether a specific nucleotide is base-paired by the introduction of a reporter molecule (generally a fluorescent label). Detection of hybridisation is fast and direct, and can be carried out on a solid support or homogeneously.

1.3.3 Fluorescent nucleotides for base mismatch discrimination

There are several general strategies for spectroscopic detection of SNPs using modified nucleotides. Modified nucleotides can be introduced at the position opposite the SNP site, so that the modified nucleotide is at the base-pair of interest (Figure 1.18, top). This can lead to a change in fluorescent properties upon hybridisation, giving either an increase in fluorescence intensity or a fluorescence quenching effect upon formation of a base-pair.
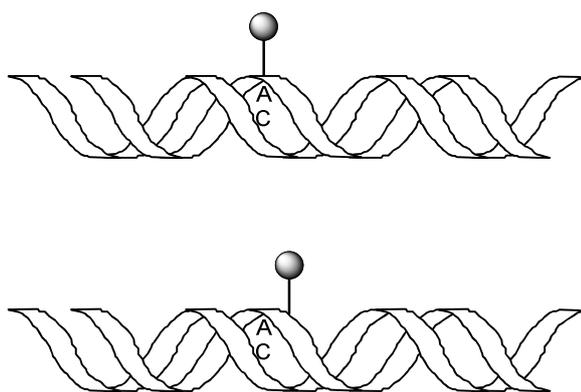
Figure 1.18 Diagram showing possible strategies for base mismatch discrimination.

An alternative strategy uses a fluorescent probe placed at an adjacent site to the base-pair of interest (Figure 1.18, bottom). It is expected that structural changes associated with a fully base-paired duplex (or the structural distortions caused by a mismatch) will influence the fluorescence intensity of the probe. Oligonucleotides which can form hairpins can also be used, where one end of the sequence is labelled with a fluorophore, and the other is labelled with a quencher. If a base-pair mismatch is present, the oligonucleotide will preferentially stay in the hairpin form, with the fluorophore quenched. If the fully complementary strand is present, the hairpin will unfold and the strands will anneal, separating the fluorophore and quencher and leading to fluorescence emission.

The Saito group has contributed extensively to the development of base-discriminating fluorescent nucleosides. The base-discriminating fluorescent nucleosides they have designed can be divided into two main categories: size expanded nucleosides and pyrene-labelled nucleosides, and include both purine and pyrimidine bases.[31] The size-expanded cytosine analogues BPP[42] and NPP[43] were reported by Saito and co-workers in 2003 (Figure 1.19). An oligonucleotide containing BPP (benzopyridopyrimidine) was hybridised with its

complementary strand. When the complementary strand contained guanosine opposite the BPP nucleotide instead of adenosine, a strong fluorescence quenching effect was observed. It was found to be suitable for identifying an A/G SNP which is known to occur in the human interferon-γ gene (IFNG).[44] Due to the very low quantum yields displayed by BPP, NPP was developed as a base-discriminating nucleotide with similar properties but slightly increased fluorescence. Both nucleotides showed intermediate fluorescence when opposite a pyrimidine base. Although they were very effective for particular sequences, these base-discriminating nucleosides were found to be unsuitable when in sequences adjacent to G/C pairs due to the fluorescence quenching observed.

Purine-based, size-expanded nucleosides were also developed by Saito and co-workers (Figure 1.19).[45] Methoxybenzoadenine (MDA) and methoxybenzoinosine (MDI) were found to be efficient for identifying pyrimidine SNPs, showing significant fluorescence quenching when an incorrect pyrimidine is opposite. However, once again the probes are quenched by G/C pairs, so are only suitable for identifying SNPs were the varying nucleotide is next to A/T pairs.
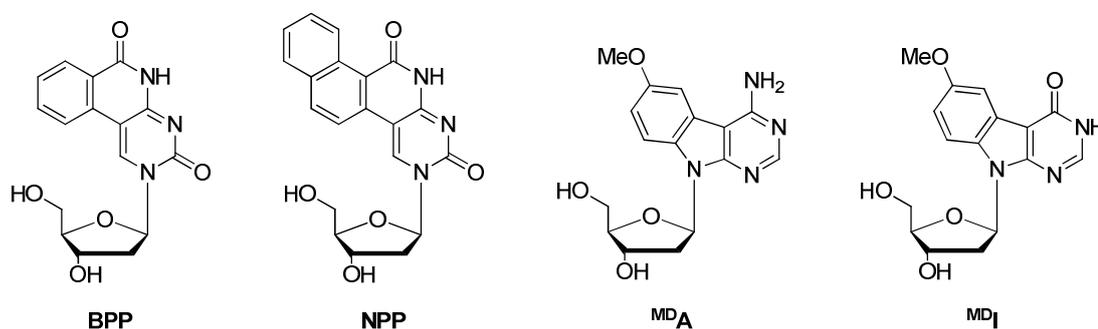
Figure 1.19 Size-expanded base-discriminating nucleosides.

Due to the sequence specificity of the size-expanded base-discriminating fluorescent nucleosides, the Saito group developed nucleosides containing pyrene fluorophores which are not in conjugation with the nucleobase. The design of these nucleosides was based on the principle that with an appropriate linker chain, the pyrene group can intercalate the DNA when the nucleoside is not base-paired, but when the correct base-pair is formed, it is suitably constrained so the fluorophore points into solution. Intercalation is known to quench pyrene fluorescence, so a quenching effect will be observed if the base-pair is

mismatched.[46] Pyrene derivatives are also known to be highly solvatochromic,[47] so changes in the polarity of the local environment will cause changes in fluorescence.

Uridine and cytidine nucleosides-bearing pyrene fluorescent labels were synthesised by Saito and co-workers (Figure 1.20).[48]

R, R' = O ^{Py}U
R = R' = H $^{Py(amino)}$U

R, R' = O ^{Py}C
R = R' = H $^{Py(amino)}$C
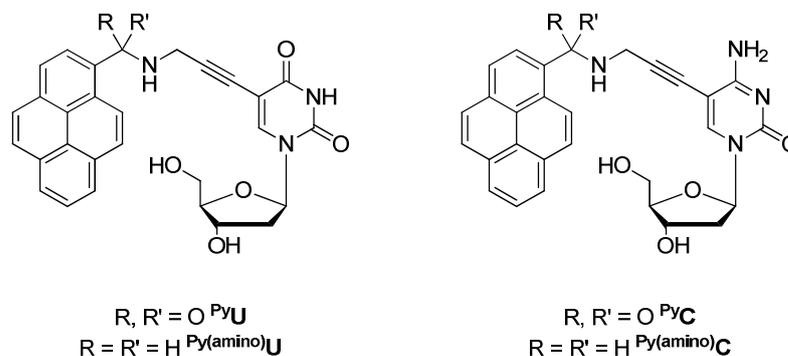
Figure 1.20 Pyrene-labelled uridine and cytidine nucleosides.

The pyrenecarboxyamide-containing nucleosides were found to be strongly fluorescent only when opposite A in oligonucleotides. In the sequence 5'-d(CGCAAT**u**TAACGC)-3' with its fully complementary strand, the quantum yield was 20.3%, a significant improvement on the size-expanded nucleosides.[31] Even in a sequence where the modified nucleotide is between cytidines, the quantum yield was still good (15.1%), unlike the size-expanded nucleotides which were strongly quenched when near G-C pairs. The labels with amino groups were only weakly fluorescent, and did not show base-discriminating properties. This supports their hypothesis that the observed solvatochromism is similar to that observed for pyrene-1-carboxaldehyde, where the solvent-dependent relative positions of the n→π* and π→π* energy levels accounts for the dramatic impact of solvent polarity on its fluorescent properties.[47] Molecular modelling (AMBER* force field in water) of the system indicated that the pyrene fluorophore was not intercalating in the DNA helix, but that in a mismatched base-pair the pyrene was bound to the DNA along the minor groove. In matched base-pairs, the pyrene fluorophore is exposed to the solvent, giving rise to the differences in fluorescence intensity.

Saito and co-workers also developed a purine analogue of the pyrene-labelled nucleosides based on 2'-deoxy-7-deazaadenosine.[49] Although the purine analogues were also synthesised with an alkyne linker, unlike the pyrimidine analogues the alkyne was removed by hydrogenation to the more flexible alkane (Figure 1.21, left). A fully complementary

duplex containing the fluorescent adenosine was found to be extremely stable – the melting temperature was found to be higher than for the unmodified duplex. The modified nucleotide was found to be highly quenched when base-paired with thymidine, but still fluorescent when base-paired with any other nucleotide. Saito and co-workers propose that the quenching effect is due to intercalation of the pyrene group, which may be possible in this case because of the more flexible propyl linker. A molecular model of a short duplex containing this nucleotide indicates that this intercalated structure is the possible cause of the base-discrimination effect (Figure 1.21, right).
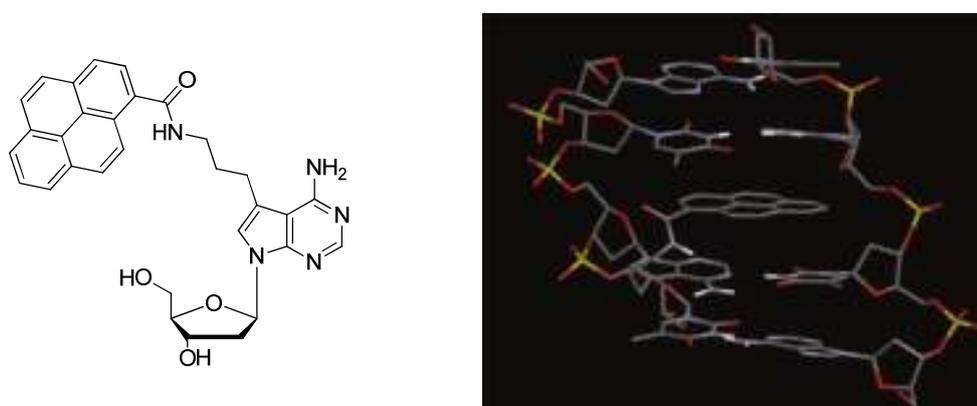
Figure 1.21 Pyrene-labelled 7-deazaadenosine analogue (left) and calculated model of matched duplex showing pyrene intercalation (right).[49]

Saito and co-workers also reported a base-discriminating fluorescent guanosine analogue,[50] similar to compounds reported by our groups.[51] Introduction of an aryl ethynyl group at C8 induces strong fluorescence in purine analogues, but also causes conformational changes (Section 1.2.1). The fluorescence brightness of the three analogues investigated by Saito and co-workers was dependent on solvent, with the dimethylamino analogue showing the highest solvent dependence. This analogue was investigated for its base-discriminating properties in two different duplexes and showed moderate base-discrimination, thought to be due to the dependency of the fluorescence emission on local polarity. Investigations in York indicate that no base-pair is formed by 8-modified purines in short duplexes.[52]
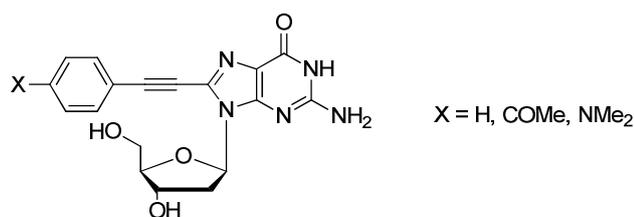
Figure 1.22 Fluorescent 8-modified-2′-deoxyadenosines investigated by Saito and co-workers.

The Tor group have also developed some fluorescent nucleosides capable of identifying base-pair mismatches. They have developed two different size-expanded analogues of uridine, based on thieno[3,4-*d*]pyrimidine[53] and quinazoline (benzopyrimidine).[54] The thieno[3,4-*d*]pyrimidine analogue (Figure 1.23, left) was phosphorylated and enzymatically incorporated into oligonucleotides by T7 RNA polymerase (RNAP). In the duplex of 5′-d(GCGCCG**u**GCA)-3′, the nucleotide was found to be strongly fluorescent when base-paired with cytidine, and showed significant quenching when opposite adenosine. However, it seems possible that if the base-discrimination is due to subtle changes in local polarity of the environment of the nucleotide rather than a specific base interaction, the identification of mismatches may be sequence dependent. The quinazoline analogue (Figure 1.23, right) has different properties, with increased fluorescence intensity when opposite a G mismatch, although again this has only been tested in one sequence. Both nucleosides were shown to form stable duplexes, with melting temperatures very similar to the unmodified duplexes.
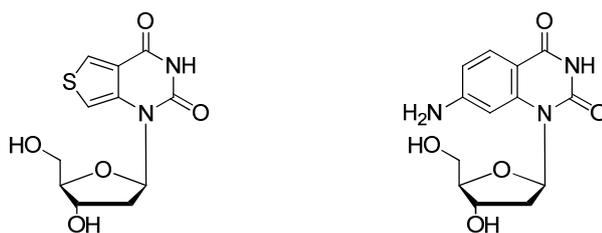
Figure 1.23 Size-expanded fluorescent base-discriminating nucleosides developed by Tor and co-workers.

Another base-discriminating, size-expanded nucleoside has been reported by Cekan and Sigurdsson.[55] A nitroxide-labelled cytidine analogue was incorporated into an oligonucleotide, and then reduced to give the amine compound which was highly fluorescent (Scheme 1.2). This fluorescent label shows moderate mismatch discrimination, with different amounts of fluorescence quenching observed for the single-stranded

oligomer depending on which base is opposite the modified base. The base-discrimination also showed sequence specificity, with better differentiation observed when the fluorescent probe was adjacent to pyrimidine bases.
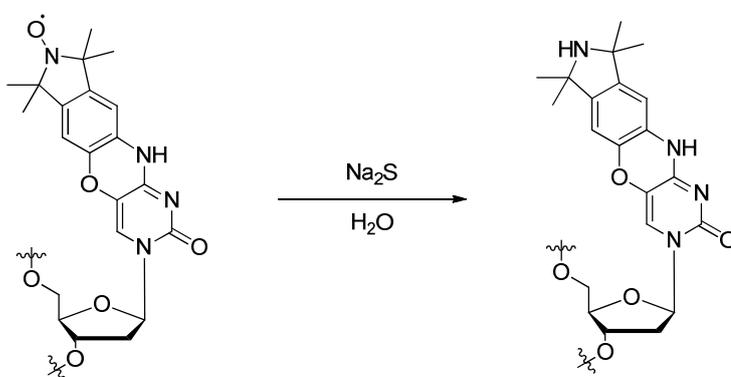
Scheme 1.2 Synthesis of fluorescent cytidine analogue by reduction of nitroxide label in oligonucleotides.

Recently, a size-expanded nucleoside based on 7-dezaadenosine was synthesised by Wilhelmsson and co-workers (Figure 1.24).[56] This simple system, consisting of a 7-phenyl-7-deazaadenosine conformationally locked by cyclisation onto the exocyclic amine, is moderately fluorescent (ϕ = 6.8% in water). The nucleoside was incorporated into short oligonucleotides with a wide range of sequences (twelve different sequences of decamers). The range of oligomers investigated allowed them to fully investigate the effect of sequence on the properties of the fluorescent nucleotide. This is more thorough than many reported examples, but is particularly important as the base-discriminating properties of some fluorescent nucleotides have been shown to be sequence dependent.[31]
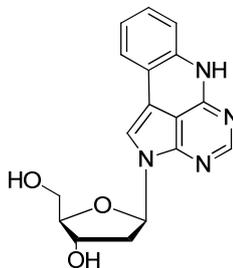
Figure 1.24 Wilhelmsson size-expanded fluorescent nucleoside.

Melting temperatures of the duplexes containing the quadracyclic adenosine are the same or higher than the melting temperatures of the unmodified duplexes, with the stabilisation observed being sequence specific. Significant destabilisation was observed for base-pair mismatches, suggesting specific thymidine base-pairing. The fluorescence quantum yields were also found to be sequence dependent, ranging from 5.8% to 0.2% for the single stranded oligonucleotides, with neighbouring pyrimidine bases causing the greatest quenching effect. In most cases, further quenching was observed upon formation of a duplex with a correctly matching base-pair. However, the presence of a base mismatch led to an increase in quantum yield being observed. This means that the modified base can be used to identify a base-pair mismatch; however the effectiveness of the base discrimination is sequence dependent.

An alternative approach has been developed jointly by the Tucker, Vyle and Bassani groups.[57] Rather than modifying the nucleoside which has a mismatch, they introduced a fluorescent probe at an adjacent site to the base pair of interest. Their method also differs because it uses a non-nucleosidic fluorescent probe (Figure 1.25, left).
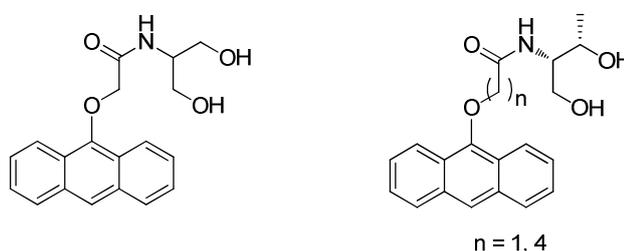
Figure 1.25 Tucker/Vyle/Bassani non-nucleosidic fluorescent probes.

The anthracene-based fluorescent probe was incorporated into an oligonucleotide which was annealed with complementary strands containing one or two mismatches at positions adjacent to the probe. Interestingly, when the modified oligomer was annealed with a fully matched complementary strand only a very small decrease in melting temperature was observed (2 °C), indicating that the anthracene may be intercalating in the duplex. This was associated with a significant drop in fluorescence intensity, also expected from intercalation of the anthracene probe (*cf.* Saito pyrene fluorescent labels). When a base mismatch (C·A) was present adjacent to the fluorescent probe, a much larger decrease in melting temperature was observed (12.5 °C), and an increase in fluorescence intensity was observed. This suggests that the anthracene is no longer intercalating in this case.

Although only one mismatch has been investigated, this probe shows potential for SNP identification.

Tucker and co-workers have also reported a similar but chiral anthracene probe, where the sugar unit is synthesised from D-threoninol and the linker is either methylene (like the previous example) or butylene (Figure 1.25, right). In this case, the fluorescent label was incorporated into duplexes and used to probe the base opposite the modification. The longer linker was found to have less impact on the melting temperature than the methylene linker, suggesting that it is better able to intercalate the duplex. The methylene-linker probe showed only very poor discrimination between different bases for both melting temperature and fluorescence methods. However, the butylene linked probe showed varying fluorescence spectra and melting temperatures for each base, suggesting that the degree of anthracene intercalation is dependent on the base opposite. Interestingly, the duplex containing this probe was also able to differentiate between cytidine and 5-methylcytidine (a common DNA methylation). Cytidine caused a small increase in fluorescence intensity compared with the single strand, while 5-methylcytidine caused an equivalent decrease in fluorescence. This highlights the high sensitivity of this fluorescent probe, although further investigation is required to determine if this is a general phenomenon.

During the preparation of this work, Hocek and co-workers reported nucleosides with fluorescent biaryl labels (Figure 1.26).[58] The modifications also contain fluorine substituents so that the labels can be used as NMR probes.

X = CH, N
R = H, NH$_2$
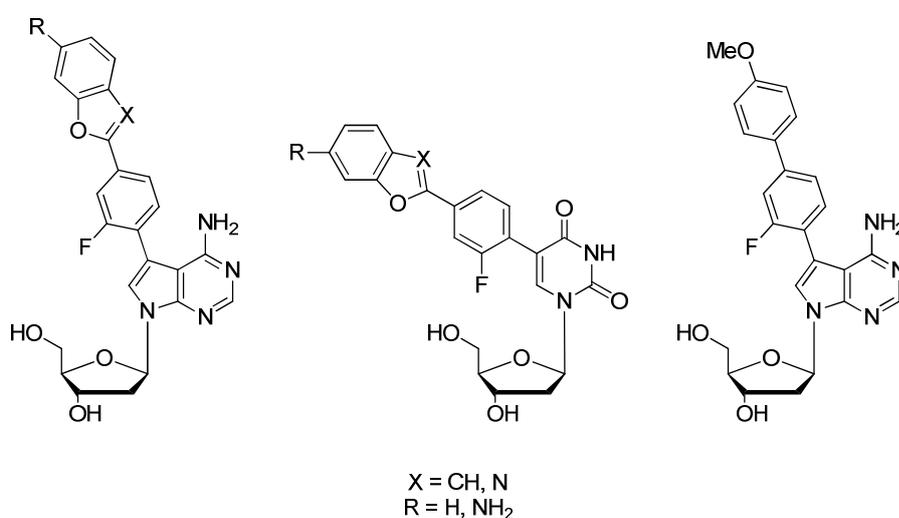
Figure 1.26 Some of the fluorine-labelled fluorescent nucleosides synthesised by Hocek and co-workers.

23

The fluorescent nucleotides were based on 2′-deoxy-7-deazaadenosine and 2′-deoxyuridine. The nucleotide triphosphates were generally found to be highly fluorescent, with quantum yields up to 28% in water. The fluorescence of the nucleotide analogues was found to be dependent on pH. The biphenyl and benzofuran-based 7-deazaadenosine analogues showed fluorescence quenching at acidic pH values, which they propose is due to protonation of the amino group. The aminobenzoxazole and benzoxazole 7-deazaadenosine analogues' fluorescence emission peaks (565 nm and 496 nm, respectively) were quenched with decreasing pH, but a second fluorescence emission (at 443 nm and 429 nm) was observed, giving a gradient of blue to yellow emission as pH varies. The uridine analogues generally showed an increase in fluorescence intensity upon increasing pH, as the nucleotide heterocycle is deprotonated. The modified nucleotide triphosphates were incorporated into oligonucleotides using primer extension (PEX), and their base mismatch discrimination investigated. The modified 7-deazaadenosine nucleotides showed some base discrimination, with moderate quenching observed when they were opposite other purines, and a larger quenching effect seen when a pyrimidine mismatch (C) was present. The uridine nucleotides showed highly sequence-dependent mismatch discrimination, with neighbouring guansoine nucleotides causing quenching of the benzofuran and benzoxazole and aminobenzoxazole-labelled uridines.

1.4   Pd-catalysed functionalisation of nucleosides and nucleotides

1.4.1   Pd-catalysed cross-coupling reactions

Palladium catalysis has become an invaluable tool for the efficient synthesis and functionalisation of complex organic compounds; moreover, palladium catalysed cross-coupling is an essential tool in the synthesis of modified nucleosides. Palladium-catalysed cross-coupling is generally used for the synthesis of sp^2-sp^2 or sp-sp^2 bonds, making it ideal for the synthesis of those fluorescent nucleosides with aryl, alkenyl or alkynyl groups directly bonded to the nucleobase heterocycle. Cross-coupling reactions require a halogenated (or pseudohalogen) species, which can be coupled to an organometallic reagent with the use of a transition metal catalyst. Typically palladium is used, but nickel and other metals are also suitable.[59] The organohalide oxidatively adds to the Pd^0 catalyst

to give a Pd^{II} species. This complex can then undergo transmetallation with the organometallic reagent, followed by reductive elimination to give the coupled product, and regenerate the Pd^0 catalyst (Scheme 1.3). Either the complex must isomerise to the *cis*-isomer for reductive elimination to occur, or ligand dissociation (to form a 14e$^-$ species) must precede reductive elimination to form the product. Often other additives, such as bases and ligands, are required.
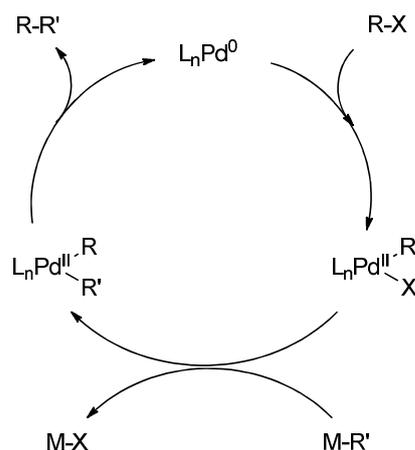
Scheme 1.3 General mechanism of palladium-catalysed cross-coupling.
The species undergoing oxidative addition, transmetallation and reductive elimination are presumed to be ligated (based on stability arguments 16-electron Pd^0 and Pd^{II} species are likely), although ligand dissociation is required to facilitate catalytic coupling, which is not specified for simplicity.

A wide variety of organometallic reagents have been used in these reactions, including organostannanes (Stille),[60] organoboronic acids (Suzuki-Miyuara),[61] Grignard reagents (Kumada),[62] organozinc compounds (Negishi)[63] and organocuprates (Sonogashira).[64] As previously mentioned, palladium-catalysed cross-coupling reactions have traditionally been used for coupling sp^2 (or sp) carbons, due to the problems associated with β-hydride elimination. However, recent catalyst systems and reaction condition development facilitate the cross-coupling of sp^3 centres.[65]

1.4.2 Pd-catalysed cross-coupling of halogenated nucleosides and nucleotides

Palladium catalysis has been extensively used in the C- and N-functionalisation of nucleosides.[66] Cross-coupling chemistry has been used for the functionalisation of all five

canonical bases as well as many unnatural or uncommon bases. The most commonly used cross-coupling reactions for nucleosides and nucleotides are Sonogashira, Suzuki and Stille cross-couplings. Traditionally, cross-coupling of nucleosides has involved protection of the sugar moiety and nucleobase before reaction to improve solubility and avoid possible side reactions. However, this is inefficient and adds extra protection and deprotection steps to the synthesis. A significant advancement was made in 1990 by Casalnuovo and Calabrese, who published a variety of cross-coupling reactions on unprotected nucleosides and nucleotides under aqueous conditions.[67] They reported conditions for both Sonogashira and Suzuki cross-couplings using the water-soluble catalyst Pd(tppms)$_3$ (TPPMS = sodium triphenylphosphine monosulfonate) in water/acetonitrile or water/methanol mixtures. These conditions could even be used for cross-coupling reactions on highly unstable nucleotide triphosphates, albeit in moderate yield (Scheme 1.4).
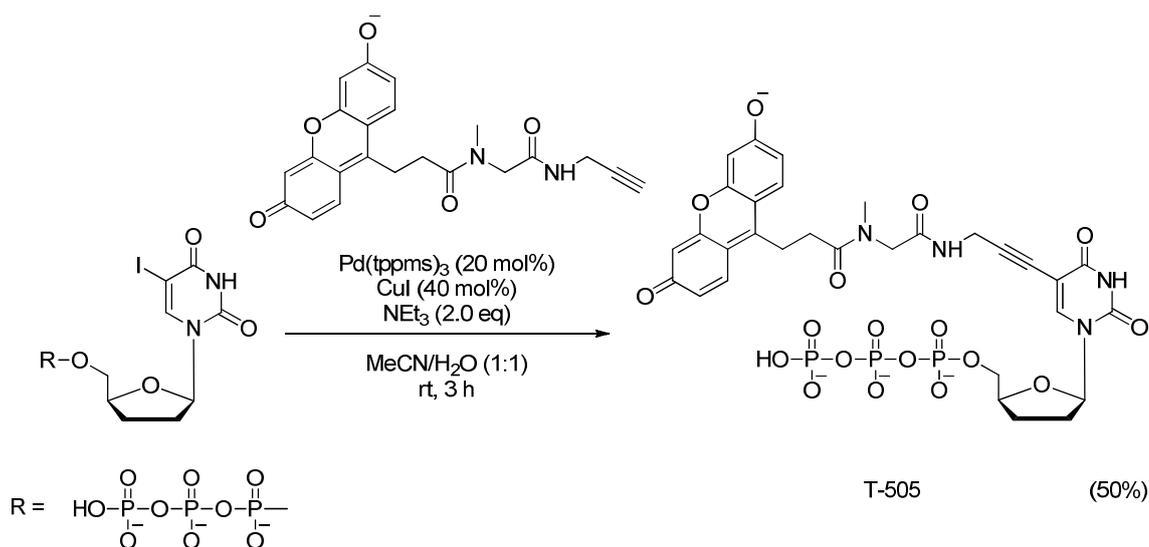
Scheme 1.4 Sonogashira cross-coupling of 5-iodo-dideoxyuridine 5'-triphosphate to give T-505.

These conditions were developed further and optimised by Shaughnessy and co-workers in 2003.[68] 8-Bromopurine and 5-iodopyrimidine nucleosides were coupled with arylboronic acids under aqueous conditions by using water soluble phosphine ligands such as TPPTS (sodium triphenylphosphine trisulfonate) to solubilise the palladium catalyst (Scheme 1.5). Although the functionalisation of 8-bromoadenosine was found to proceed readily, guanosine was found to be a more challenging substrate due to the readily ionisable amine group and multiple binding modes to palladium. Shaughnessy and Western have

subsequently shown the inhibitory effect guanosine may have on palladium catalysed cross-coupling reactions.[69] They also found that TXPTS (trisodium tri(4,6-dimethyl-3-sulfonatophenyl)phosphine) was a more effective ligand for nucleoside cross-coupling than TPPTS, giving quantitative yields at room temperature in 30 min.
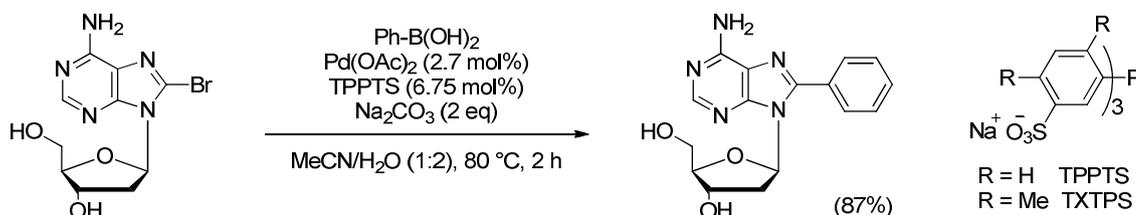
Scheme 1.5 Aqueous Suzuki-Miyaura cross-coupling of 8-Br-adenosine with phenylboronic acid.

Collier and Wagner improved the synthesis of 8-arylguanosine mono- and triphosphates by using a pyridine additive and acetonitrile solvent in the phosphorylation of 8-bromoguanosine.[70] 8-Bromo-ATP and AMP could then be used in cross-coupling reactions using similar conditions to Shaughnessy and co-workers. Their process is an efficient two step synthesis of 8-substituted guanosine triphosphates or nucleotides from 8-bromoguanosine. They also applied Suzuki-Miyaura cross-coupling chemistry to the synthesis of modified NAD$^+$ analogues.[71] NAD$^+$ was brominated and then directly cross-coupled *in situ* to give 8-aryl-NAD$^+$ derivatives (Scheme 1.6)
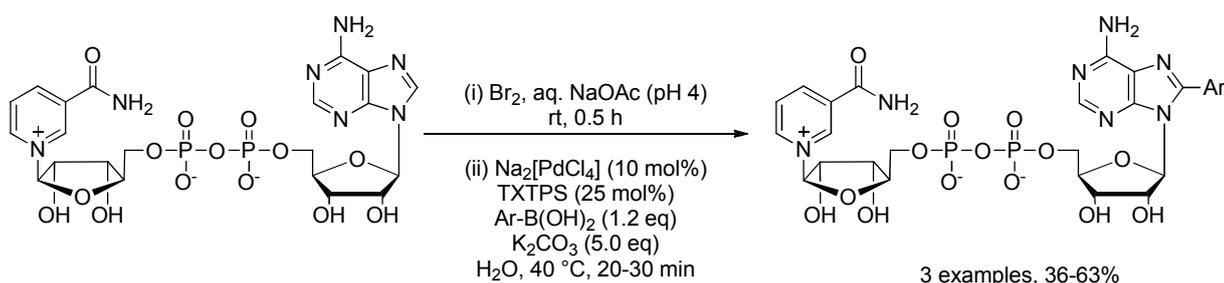
Scheme 1.6 One-pot synthesis of 8-aryl-NAD$^+$ from NAD$^+$.

The chemistry developed by Casalnuovo and Calabrese, and by Shaughnessy and co-workers, has been applied to a range of nucleoside and nucleotide substrates, to introduce a wide array of substituents. Some of these nucleoside analogues will be discussed in later

sections, but notable contributions have come from the Hocek group, who in particular have investigated the metal:ligand ratio and developed microwave conditions for this chemistry.[72] They have also optimised the conditions for the Suzuki cross-coupling of halogenated nucleotide triphosphates, for an efficient and significantly less divergent route to functionalised nucleotides (Scheme 1.7).[73]
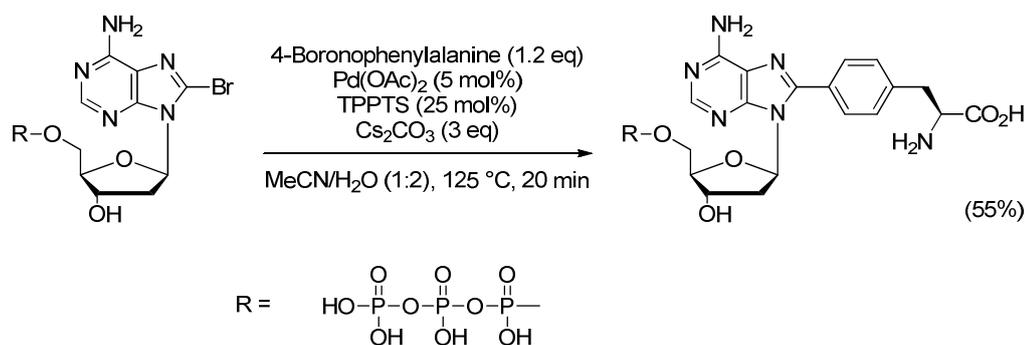
Scheme 1.7 An example of cross-coupling on an unprotected nucleotide triphosphate using Hocek conditions.

One of the most remarkable applications of Shaughnessy's conditions reported to date is the cross-coupling of a brominated nucleobase incorporated into an oligonucleotide. Manderville and co-workers incorporated 8-bromoguanosine into short oligonucleotides (up to 15mers), which were then subjected to Pd-catalysed cross-coupling using conditions based on those reported by Shaughnessy (Scheme 1.8).[74] A large excess of arylboronic acid was required to give the arylated-oligonucleotides, which were isolated in good yields (45%).
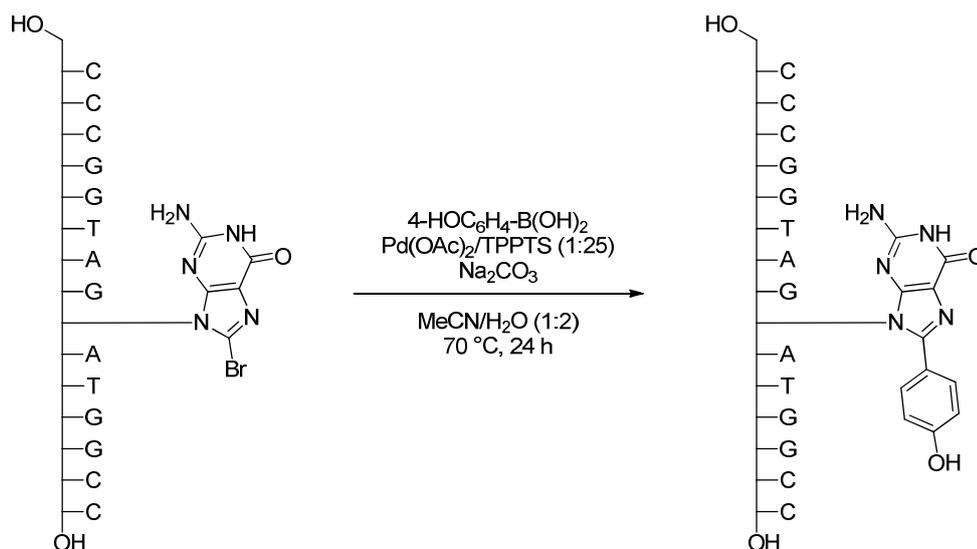
Scheme 1.8 Manderville conditions for cross-coupling of brominated oligonucleotides.

A few examples of Sonogashira cross-coupling of unprotected 8-bromoadenosine with terminal acetylenes were reported by Cristalli and co-workers in 2001, although no characterisation of the products was given.[75] Liebscher and co-workers also synthesised a variety of lipophilic 8-substituted adenosines by Sonogashira cross-coupling, however synthesis of the 8-substituted guanosine analogues required protection of the sugar hydroxyls and the exocyclic amine.[76]

Previously in the Baumann and Fairlamb groups, Firth investigated the Sonogashira cross-coupling of unprotected purine nucleosides, including the first Sonogashira cross-coupling of unprotected 8-bromoguanosine.[51] A library of 8-substituted adenosines and guanosines was synthesised by Sonogashira cross-coupling, including thienylacetylene- and ferrocenylacetylene-substituted products (Scheme 1.9). The effect of the Pd:Cu ratio in the Sonogashira cross-coupling of guanosines was investigated. The optimal conditions were found to require 2 mol% Cu and 1 mol% Pd.
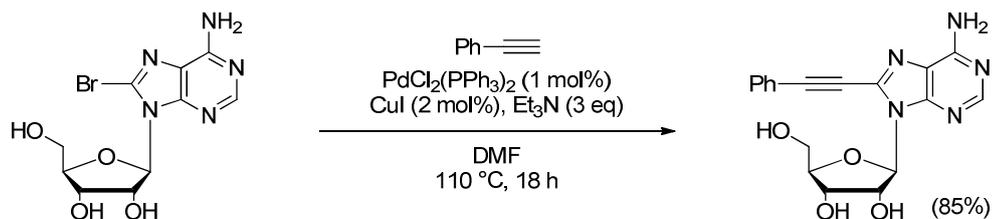
Scheme 1.9 Sonogashira cross-coupling of adenosine with phenylacetylene.

Similar aqueous conditions to those used by Shaughnessy for Suzuki-Miyaura cross-couplings have also been applied to Sonogashira cross-couplings. Some of the first examples of this type were published by Hocek and co-workers for the cross-coupling of halogenated nucleotide triphosphates (Scheme 1.10).[77] Shaughnessy and co-workers subsequently optimised the conditions for the unprotected 8-bromopurine and 5-iodopyrimidine nucleosides, finding that, as for Suzuki cross-couplings, TXTPS was the most efficient water-soluble phosphine ligand.[78]

(cytosin-5-yl)acetylene (2.2 eq)
Pd(OAc)$_2$ (4 mol%)
TPPTS (20 mol%)
CuI (10 mol%)
iPr$_2$EtN (11.2 eq)

MeCN/H$_2$O (1:2), 80 °C, 1 h

(17%)
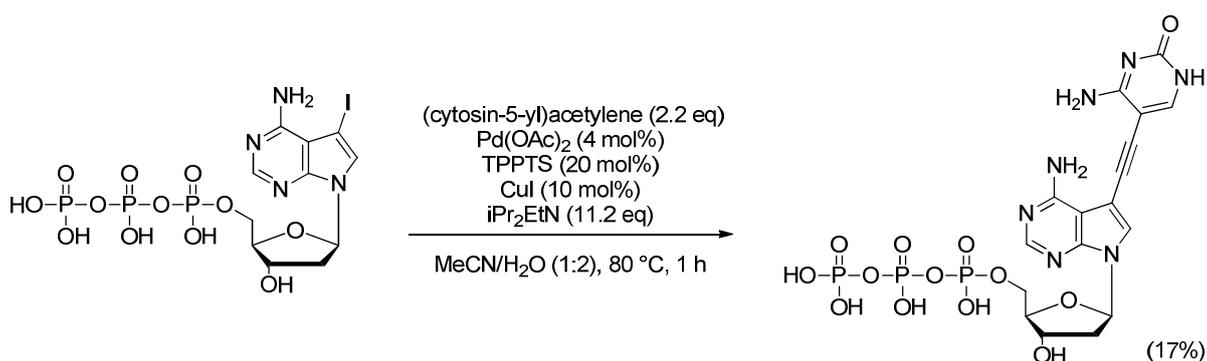
Scheme 1.10 Aqueous Sonogashira cross-coupling of halogenated nucleotide triphosphates.

Another earlier application of Pd-catalysed cross-coupling to the modification of oligonucleotides was published by Richert and co-workers in 2004.[79] Their application of Sonogashira cross-coupling to halogenated oligonucleotides is similar to that published later by Manderville and co-workers, but Richert's method differs in that the oligonucleotides are protected and still bound to the solid-phase support used for synthesis of the oligonucleotides. This is advantageous as it makes purification of the products much easier. The conditions are also quite different from Manderville's; rather than using an aqueous Pd-catalyst such as Pd(OAc)$_2$ with TPPTS, more traditional Sonogashira cross-coupling conditions are used (Scheme 1.11).
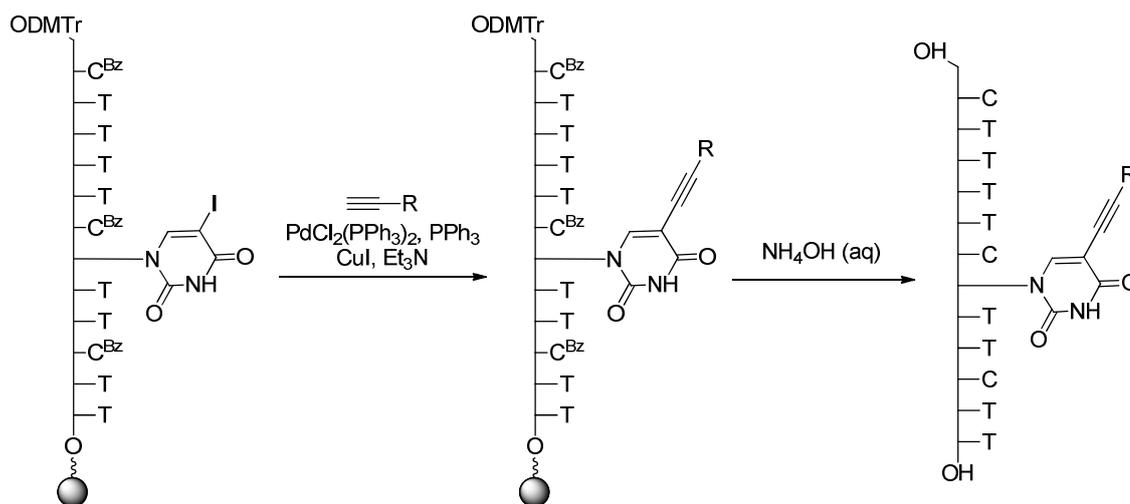
Scheme 1.11 Richert Sonogashira alkynylation of 5-I-U in an oligonucleotide bound to a solid support.

1.4.3   Pd-catalysed C-H functionalisation of nucleosides

Direct arylation is the coupling of an aryl halide with a C-H bond, often a heterocycle (Scheme 1.12). Direct arylation methodology has become more widely applicable to a range of heterocycles in recent years.[80] It is advantageous over traditional cross-coupling routes as it avoids the pre-functionalisation of the heterocycle, does not require the use of expensive (arylboronic acids) or toxic (organotin) coupling partners, and often increases atom-efficiency.

$$\text{Ar}-\text{H} \ + \ \text{Ar'}-\text{X} \ \xrightarrow[\text{base}]{\text{cat.}} \ \text{Ar}-\text{Ar'} \ + \ \text{HX.base}$$
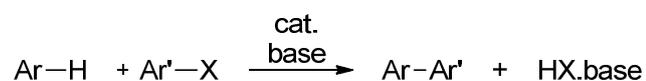
Scheme 1.12 Direct arylation methodology.

The direct arylation of *N*-benzylpurines was reported by Hocek and co-workers in 2006. 2,6-Dichloropurines were substituted at C8-H by cross-coupling with aryl halides using a Pd catalyst and excess copper(I) iodide and base.[81] Hocek developed the direct arylation of unprotected adenosine nucleosides simultaneously with the Fairlamb group.[82,83] Both processes employ a palladium catalyst and excess copper(I) iodide. The Fairlamb conditions require an excess of caesium carbonate (Scheme 1.13), whereas Hocek uses

piperidine as a base. In both cases, diarylated products, where both the 8-position and the exocyclic amine group were arylated, were by-products, however this was more problematic under the Hocek conditions (Fairlamb – 4 %, Hocek – 18 %). The higher temperatures used by Hocek also led to increased deglycosylation, and therefore lower yields. The Fairlamb group also investigated the reaction mechanism. Although Pd[0] nanoparticles were formed under the reaction conditions, a negative mercury drop test suggested the reaction to be homogenous. However, the efficacy of this test as a probe for the homogeneity of a reaction has not been fully determined, particularly in very polar solvents such as DMF.[84,85] The role of copper(I) iodide was not fully understood despite these investigations. A Sonogashira-like mechanism was proposed, however this does not account for the requirement for three equivalents of copper iodide (Scheme 1.14). Further investigation into this reaction within the Fairlamb group has suggested that coordination of the nucleobase to Cu[I] plays a key role in this reaction, and may account for the need for an excess of copper iodide.[86]
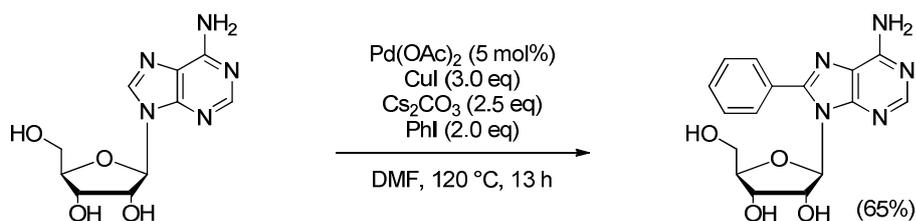
Scheme 1.13 Direct arylation of adenosine under Fairlamb conditions.
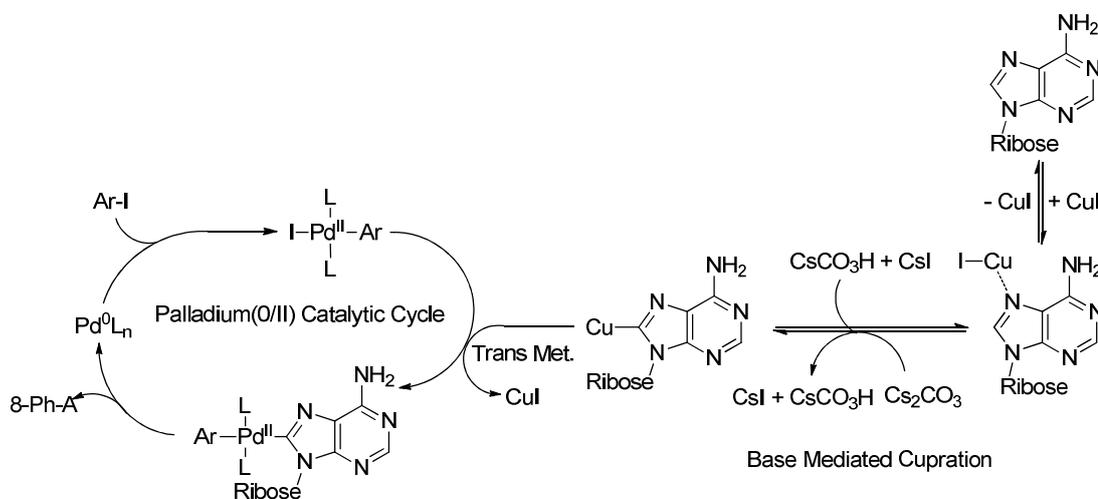
Scheme 1.14 Proposed mechanism for Pd/Cu mediated direct arylation of adenine nucleosides.

Some insight into the mechanism may be gained from looking at simpler systems previously reported in the literature. Pd/Cu systems were first developed by Miura and co-workers in the 1990s for the direct arylation of heterocycles.[87] They found that the selectivity of Pd-catalysed arylations was affected by the presence of Cu[I] salts. When Pd(OAc)$_2$ was used as the sole catalyst, the arylation of N-methylimidazole with iodobenzene was selective for C-5, however when a stoichiometric amount of CuI was added, the selectivity was pushed towards C-2 arylation (Scheme 1.15)

Pd (10 mol%)	54%	0	24%
Pd (10 mol%)/Cu (2.0 eq)	0	37%	40%
Cu (2.0 eq)	0	37%	0
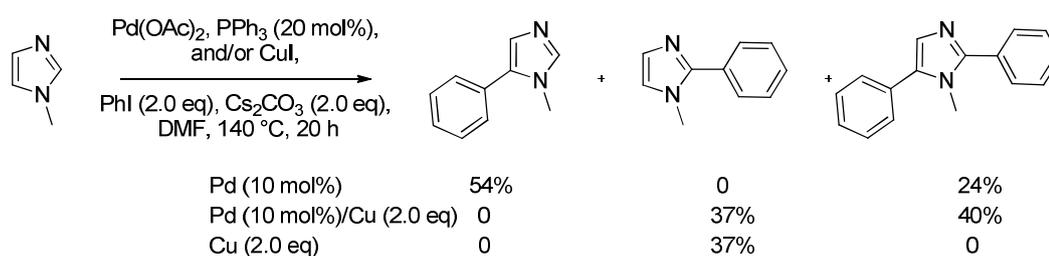
Scheme 1.15 Effect of CuI on direct arylation of N-methylimidazole.

Bellina, Rossi and co-workers reported an investigation into the effects of additives on the palladium-catalysed direct arylation of imidazoles, which led to the optimisation of conditions for selective C-2 arylation (Scheme 1.16).[88] They proposed that the switch in selectivity is due to a change in mechanism; arylation at C-5 occurs *via* electrophilic aromatic substitution by the 'ArPd^{II}X' species, whereas arylation at C-2 occurs *via* cupration at C-2, followed by transmetallation to Pd.

13 examples, 52-94%
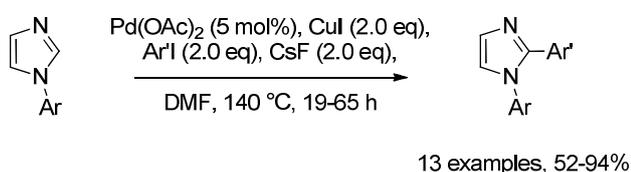
Scheme 1.16 Direct arylation of imidazole at C2.

The Fairlamb group then developed the direct arylation of 2'-deoxyadenosine.[86] This is a more challenging substrate due to the increased instability of the glycosyl bond (10^2 times more unstable than adenosine).[89] Lowering the reaction temperature from 120 to 80 °C was required to prevent deglycosylation (Scheme 1.17). It was found that the reaction at

this temperature gave unreliable yields depending on how the solvent was purified. On further investigation, it was found that this was due to decomposition products of DMF which were not formed at the lower temperature. DMF can decompose on heating to form dimethylamine and carbon monoxide.[90] After screening of a variety of amines, it was found that sub-stoichiometric piperidine (40 mol%) was an adequate 'dimethylamine mimic', giving good, repeatable yields.[86] It may aid the reduction of $Pd(OAc)_2$ (*i.e.* the secondary amine lowers the reduction potential at Pd^{II}), *via* $[Pd(HNC_5H_{10})_2(OAc)_2]$, to Pd^0 under these more mild reaction conditions.
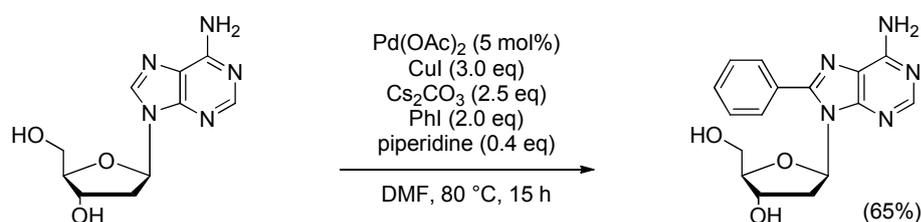
Scheme 1.17 Direct arylation of 2'-deoxyadenosine.

Nucleosides with extended aryl systems can be synthesised by carrying out the direct arylation reaction on 2'-deoxyadenosine with 3- or 4-chloro- or bromo- iodobenzene, selectively functionalising the iodide over the other halogens.[91] Further Suzuki-Miyaura cross-coupling of the second halogen gives biphenyl products with extended π-systems (Scheme 1.18). If a third halogen is incorporated during the Suzuki cross-coupling step, a third aryl group can be added in a second Suzuki arylation step.
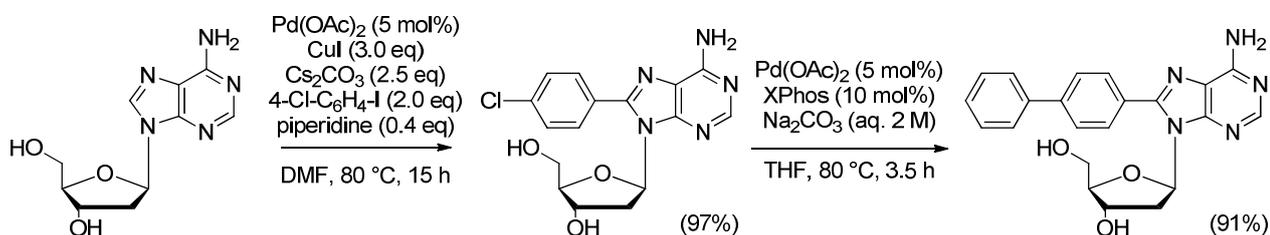
Scheme 1.18 Synthesis of 8-biaryl-2'-deoxyadenosines by sequential direct arylation and Suzuki-Miyaura cross-coupling.

1.5 Aims and Objectives

The central aim of this project is the design and synthesis of a novel class of base-modified fluorescent nucleosides and nucleotides. Whilst there are many commercially available fluorescent nucleosides, there is a scarcity of fluorescently labelled purine analogues with biochemical properties similar to the natural nucleosides. Fluorescently labelled purine nucleosides and nucleotides generally contain fluorescent groups with long linker groups, such that the label is distant from the nucleobase. This is useful for some applications, but often it is preferable to have a fluorescent label which is less bulky and can act as a reporter of the local chemical environment.

The development of fluorescent adenosine analogues, with properties similar to the natural nucleoside, which can act as probes of local environment, is highly desirable. There are many potential applications for fluorescently labelled nucleosides and nucleotides. This project will focus on their applications as base-discriminating fluorescent nucleotides for identification of single-nucleotide polymorphisms. Currently reported base-discriminating fluorescent nucleotides often have limitations, such as limited sequence specificity or quenching by neighbouring bases. A fluorescent nucleotide with the fluorophore in conjugation with the nucleobase should have fluorescence which is sensitive to base-pair formation. Crucially, the base-discriminating properties of the nucleotide will be less affected by subtle geometry changes which are imparted by sequence variation.

Modification at C7 of 7-deazaadenosines has been shown to be a useful way of introducing fluorescent labels to adenosine nucleotides. 7-Modified-7-deazaadenosines exhibit high duplex stability with little duplex distortion, and high DNA melting temperatures. Alkynyl-substituted 7-deazaadenosines, where the conjugated alkynyl group acts as the fluorophore (Panel A, Figure 1.27) generally possess low quantum yields and/or absorbance maxima overlapping with the intrinsic absorbances of other biomolecules.

Previous experience with 8-modified-purines has shown that the introduction of relatively small substituents (such as phenyl and phenylethynyl) groups at C8 is adequate to introduce significant fluorescence. In these systems, conjugation between the π electrons in the C8 substituent group and the purine nucleobase occurs readily. However, the π electron systems of chemical substituents at C7 are less likely to conjugate with the

nucleobase. This is the likely explanation for the aforementioned poor fluorescent properties of phenyl- and phenylethynyl-substituted 7-deazapurines. For this reason, a novel organic rigid rod system based on a phenylethynylbenzene motif, or a phenylethynyl group linked to a terminal hetero-aromatic group, will be synthesised to induce adequate fluorescence (Panel B, Figure 1.27). The purine nucleobase at one end of the system will likely influence its intrinsic fluorescent properties. Previous experience in the 8-modified series has shown that substitution of the terminal phenyl group (located at the other end of the rigid rod system) with electron withdrawing or donating groups can then be used to tune the fluorescence properties of the analogue. Therefore a series of 7-deazapurines containing electron donating and electron withdrawing rigid-rod fluorophores at C7 will be synthesised.

Panel A:
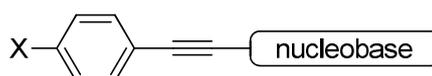
Panel B:
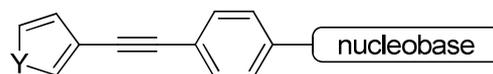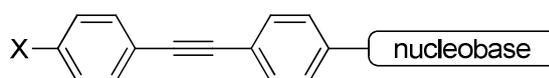
Figure 1.27 Rigid-rod fluorescent labels for 7-deazapurine nucleosides.

A related area of interest is the development of methodology for the synthesis of C-modified nucleosides. This work was initiated previously with the optimisation of Sonogashira coupling conditions for application to purine nucleosides,[51] and then with the development of C-H functionalisation methodology for direct arylation at C8 of adenosine,[86] which employed a Pd catalyst and stoichiometric Cu additive. C-H functionalisation is advantageous as it avoids the need for pre-functionalisation (halogenation) of the nucleoside substrates. Direct arylation potentially allows for functionalisation of the unmodified nucleoside in a biological setting, *e.g.* for *in situ* labelling of nucleic acids. Because of the inaccessibility of the C8-H position of adenosine when the nucleoside is base-paired, only exposed adenine residues (such as in mismatches

and hairpins) could be functionalised. However, for this to be a practical and compatible methodology, milder conditions must first be developed. Understanding the reaction mechanism is central to improving this catalytic methodology.

In summary, the aims of the project are:

- Design and synthesis of a novel class of 7-deazaadenine nucleosides with fluorescent substituents at C7.
- Incorporation of 7-deazaadenine nucleosides with appropriate photophysical properties into oligonucleotides and determine their suitability as base discriminating fluorescent probes.
- Improvement and extension of C-H functionalisation methodologies for accessing fluorescent nucleosides and nucleic acids.

The following objectives will be key to achieving these aims:

- Characterisation of the fluorescent nucleosides by UV-Vis spectroscopy, including solvatochromic studies.
- Characterisation of the fluorescent nucleosides by fluorescence spectroscopy, including both steady-state and time-resolved measurements.
- Incorporation of the fluorescent nucleosides into oligonucleotides using solid-phase phosphoramidite chemistry.
- Determine if incorporation of the fluorescent nucleoside alters the secondary structure of the oligonucleotide using circular dichroism spectroscopy.
- Assessment of the base-discriminating properties of the fluorescent nucleotide within the oligonucleotides by examination of the effect of base-pair mismatches on secondary structure, duplex stability and the fluorescence properties of the labelled nucleotides.
- Further investigation into the mechanism of the Pd-catalysed direct arylation of 2′-deoxyadenosine, in particular the role of CuI.

2 Synthesis of nucleoside analogues

2.1 Introduction

In the 1960s, Davoll synthesised a series of pyrrolo[2,3-*d*]pyrimidines (7-deazapurines) using a relatively simple procedure, making 7-deazapurine nucleosides more accessible.[92] The total synthesis of tubercidin and related nucleosides and deoxyribonucleosides has been achieved by stereospecific nucleobase anion glycosylation of 6-chloro-7-deazapurine, **3** (Scheme 2.1).[93,94] The stereoselectivity in this reaction is controlled by the temperature of synthesis of the 1-chloro-ribofuranose, producing either the thermodynamic (**2β**) or kinetic (**2α**) product.
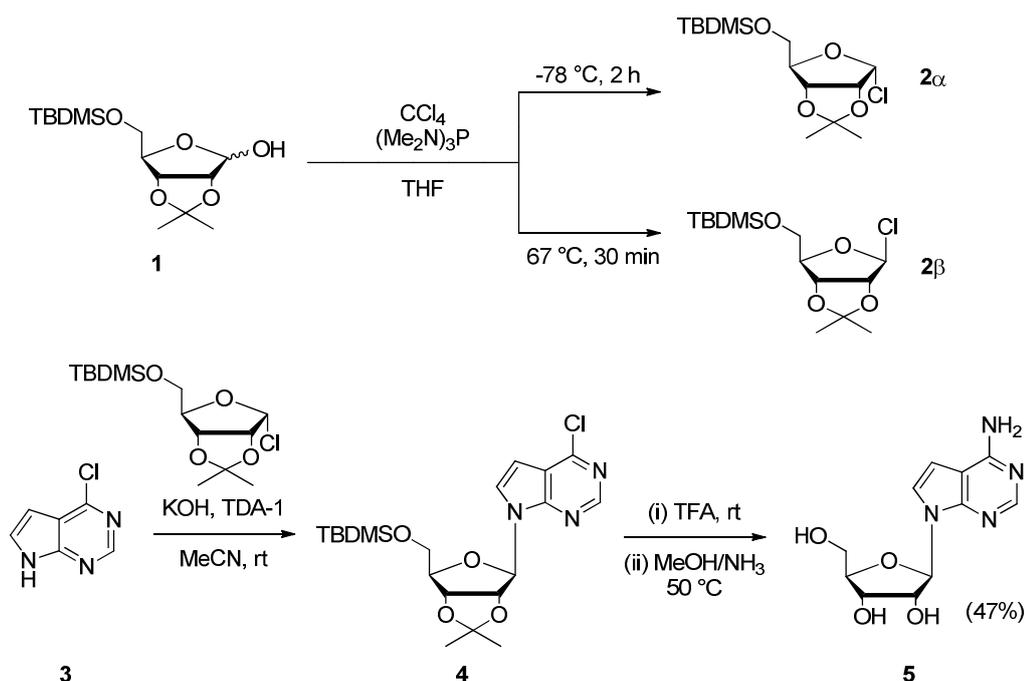
Scheme 2.1 Synthesis of tubercidin from 6-chloro-7-deazapurine.

A similar route can also be used to synthesise the 2'-deoxyribose analogue,[95] using Hoffer's chlorosugar, **6** (Scheme 2.2).[96] Hoffer's chlorosugar is often used in the synthesis of 2-deoxyribose-containing compounds as it is simple to prepare, and a single anomer of the sugar can be produced with >99% selectivity. The selectivity observed with **6** suggests that the reaction proceeds *via* an S$_N$2 mechanism (**I**, Figure 2.1), rather than *via* dissociation of

the chloride to form an oxonium intermediate (**II**, Figure 2.1). The formation of only β-nucleosides from reactions with tri-*O*-benzyl-α-D-arabinofuranosyl chloride showed that steric effects do not play a role in this reaction, and the selectivity was rationalised by the presence of a close ion-pair intermediate (**III**, Figure 2.1).[97]
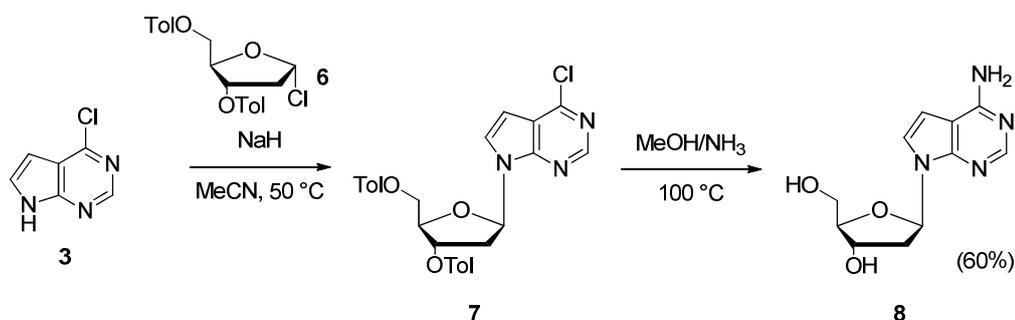
Scheme 2.2 Synthesis of 2′-deoxytubercidin from Hoffer's chlorosugar.
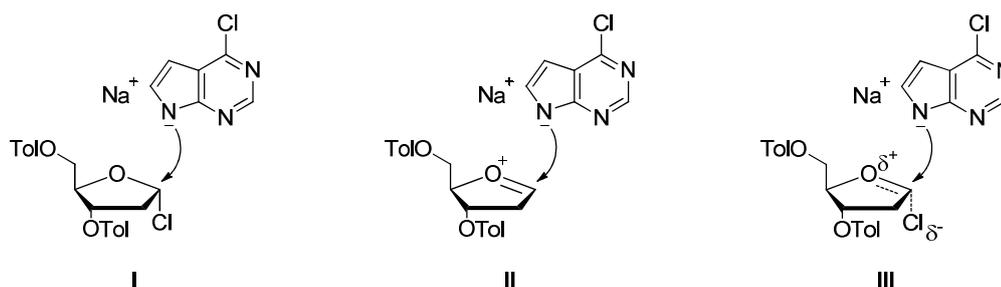
Figure 2.1 Possible intermediates in the glycosylation of 6-chloro-7-deazapurine.

Glycosylation of nucleobases under Vorbrüggen conditions (also known as the silyl-Hilbert-Johnson reaction) is a versatile method for accessing new nucleoside analogues in a one-pot process.[98,99] Vorbrüggen glycosylation requires silylation of the nucleobase NH and activation of a sugar protected with acyl groups and activated by addition of a Lewis acid (Scheme 2.3).[100] The α/β selectivity is determined by the configuration at the 2-position of the sugar by neighbouring group participation, *i.e.* where the 2-hydroxyl is protected as an acyl group. In most recent examples, the nucleobase is silylated using *N,O*-bis(trimethylsilyl)acetamide (BSA), and trimethylsilyl triflate is used as the Lewis acid.[101] The reaction is therefore catalytic in BSA and TMSOTf, although stoichiometric amounts are often used to achieve shorter reaction times.
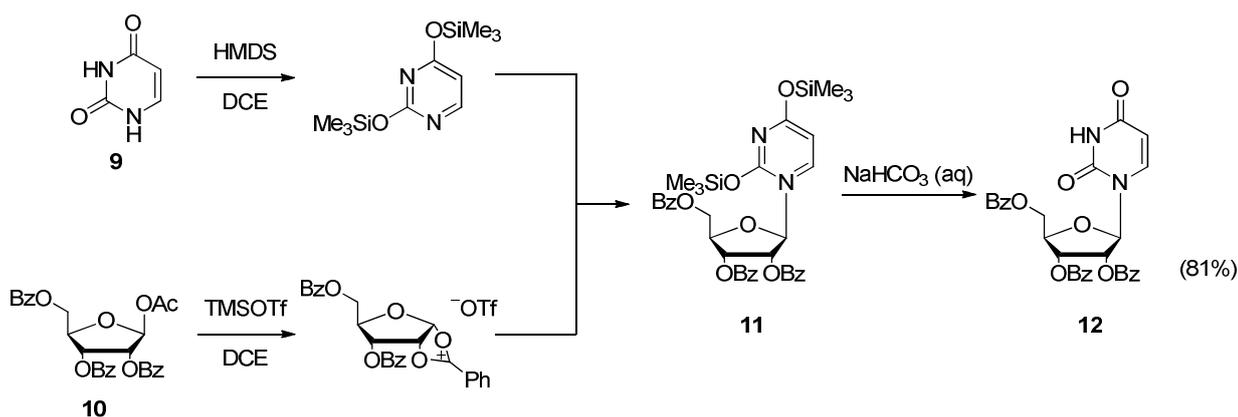
Scheme 2.3 Vorbrüggen glycosylation of uracil.

The reaction has a wide substrate scope and can be used for the glycosylation of both purine and pyrimidine bases. However, the pyrrole nitrogen of 7-deazapurines is inert under these conditions. Instead, 7-deazapurines will either react at C7 or at one of the pyrimidine nitrogens, or give no product.[102] Surprisingly, the 7-halogenated analogues readily undergo glycosylation under identical conditions. This may be due to the electronic influence of the halogen or the increased thermal stability of the product at higher reaction temperatures (Scheme 2.4).[103] The reactivity of *N*-9 may also be increased by the reactive C-7 position being blocked in these 7-substituted derivatives.
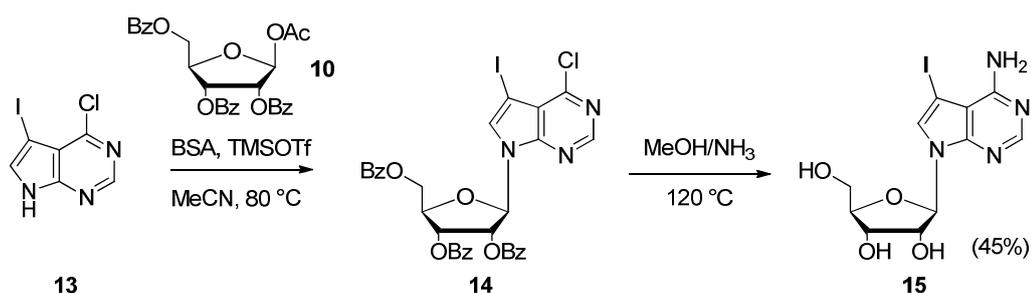
Scheme 2.4 Glycosylation of 7-substituted-7-deazapurines.

Glycosylation of 7-halogenated-7-deazapurines is useful because it facilitates access to a wide range of 7-substituted derivatives. The halogen substituents allow for further substitution, either by traditional organic chemistry, or by palladium-catalysed cross-coupling reactions (See Section 1.4.1).

2.2 Synthesis of 7-iodo-7-deazaadenosine – a versatile starting material for synthetic elaboration

A three-step synthesis of 7-iodo-7-deazaadenosine from the commercially available reagent 6-chloro-7-deazapurine has been previously published.[103] Firstly, the purine heterocycle is iodinated, then the protected ribose sugar can be attached using Vorbrüggen glycosylation, and finally a global deprotection reaction with ammonia gives the deprotected nucleoside and installs the exocyclic amine.[103] This is the synthetic route followed in this work, although optimisation was necessary (Scheme 2.5)
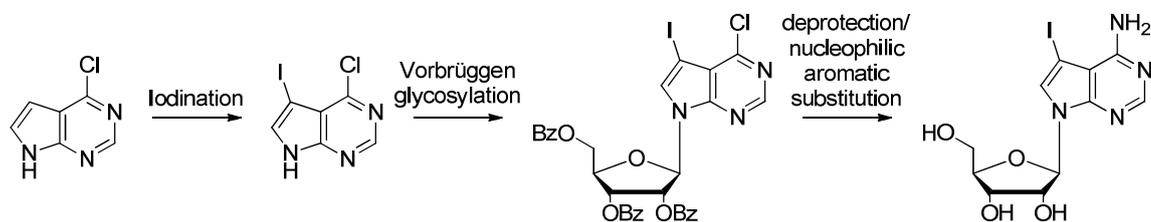
Scheme 2.5 Synthetic route to 7-iodo-7-deazaadenosine.

The synthesis of 6-chloro-7-iodo-7-deazapurine, **13**, reported by Seela and co-workers uses *N*-iodosuccinimide (NIS) as the iodinating agent, followed by recrystallisation from methanol to give the product.[27] However, it was found that several recrystallisations were required to achieve high purity and good yields (>70%) and a by-product was observed in 5-10% yield, resulting from nucleophilic aromatic substitution of the chlorine substituent by methanol, to give 6-methoxy-7-iodo-7-deazapurine. As this by-product is undesired and difficult to separate from the product, a new method for the iodination reaction was required. Iodine in DMF has been used previously for the iodination of indoles.[104] This procedure was applied to 6-chloro-7-deazapurine, and the iodinated product was isolated in 74% yield (Scheme 2.6). This procedure has been used to carry out the synthesis of **13** on a multigram scale (17 mmol).
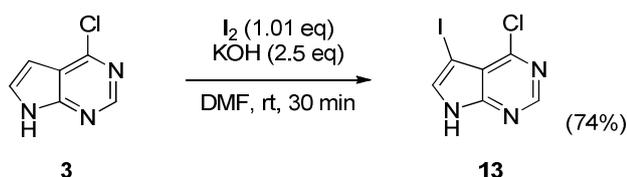
Scheme 2.6 Synthesis of 6-chloro-7-iodo-7-deazapurine.

Glycosylation could then be carried out under Vorbrüggen conditions, using TMSOTf and BSA (Scheme 2.7). The glycosylation reaction was carried out according to the procedure reported by Seela *et al.*[103] With these conditions, the glycosylated product was isolated in 64 % yield. Optimisation of this step was undertaken in an attempt to reduce the quantity of BSA and TMSOTf required, and to increase the yield (Table 2.1).
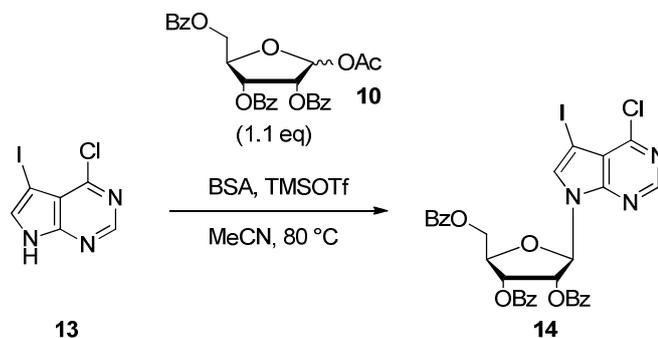
Scheme 2.7 Glycosylation of 6-chloro-7-iodo-7-deazapurine.

Table 2.1 Effect of reagent stoichiometry on glycosylation reaction.

BSA	TMSOTf	Yield^a /%
1.2 eq	1.1 eq	64 (59 lit.[103])
0.5 eq	0.5 eq	71
0.1 eq	0.1 eq	33^b

^{a} After chromatography on silica gel. ^{b} Isolated as 1:1 mixture of protected sugar and product.

It was found that a sub-stoichiometric amount of both reagents marginally increased yield (71%), but decreasing the ratio to 0.1 equivalents resulted in a significant drop in yield (33%).

As previously mentioned, the deprotection/nucleophilic aromatic substitution step reported by Seela *et al.* was carried out in an autoclave with methanolic ammonia at 120 °C. The reaction was attempted at 80 °C in a sealed tube and gave 45% yield, along with some of the 6-methoxy analogue (formed via competing reaction of the chloro compound with methanol). Due to the practical difficulties encountered repeating these conditions without access to an ammonia-resistant autoclave, milder conditions were required. As previously reported, in cases where substitution by methanol is problematic, aqueous ammonia/dioxane at 60 °C can be used as an alternative.[103] This approach proved more successful, eliminating the problems encountered using ammonia under high pressure. After 3 days, the product was isolated in 87% yield (Scheme 2.8). The 7-iodo nucleoside product was formed as a single diastereomer, with NMR characterisation consistent with the literature.
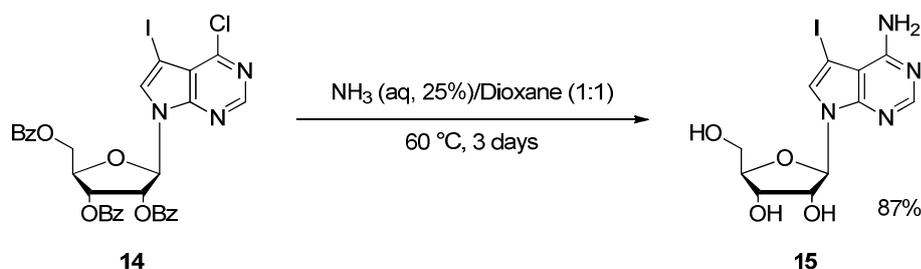
Scheme 2.8 Deprotection/nucleophilic aromatic substitution.

2.3 Synthesis of 7-modified fluorescent nucleosides

2.3.1 Initial cross-coupling reactions

As discussed previously, palladium-catalysed cross-coupling is a powerful tool for the synthesis of arylated and alkynylated nucleosides from the corresponding halogenated substrates. Sonogashira cross-coupling of **15** with phenylacetylene, using the conditions developed by Firth *et al.*[51], resulted in formation of the product in 65% yield. However the product was contaminated with a mixture of unknown by-products which could not be separated by column chromatography on silica gel. Lowering the reaction temperature from 80 °C to room temperature (*ca.* 13-23 °C) gave a cleaner reaction, giving **16** in 83% yield (Scheme 2.9).
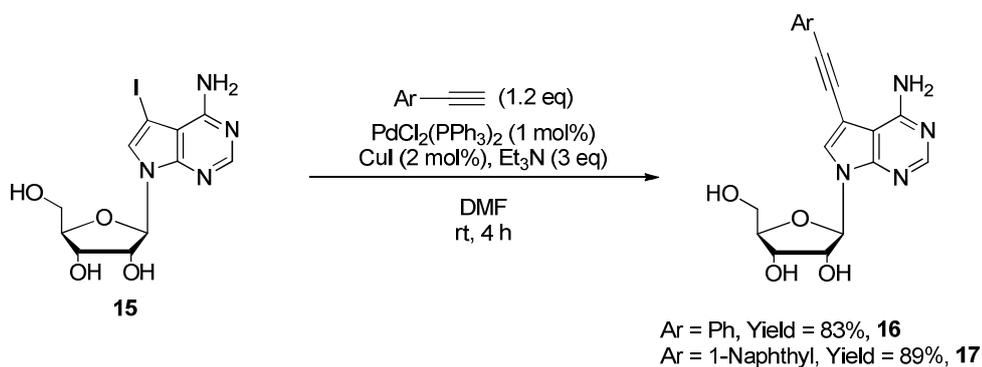
Scheme 2.9 Sonogashira cross-coupling of 7-iodo-7-deazaadenosine with phenylacetylene.

7-(1-Naphthyl)-7-deazaadenosine, **17** was also synthesised by this method. The product was isolated as a yellow solid in 89% yield following chromatography on silica gel. 7-Phenyl-7-deazaadenosine was synthesised by Suzuki-Miyaura cross-coupling, using the conditions developed by Hocek and co-workers.[105] The reaction was carried out under aqueous conditions, using TPPTS as the water soluble ligand, and **18** was isolated in 72% yield (Scheme 3.5).
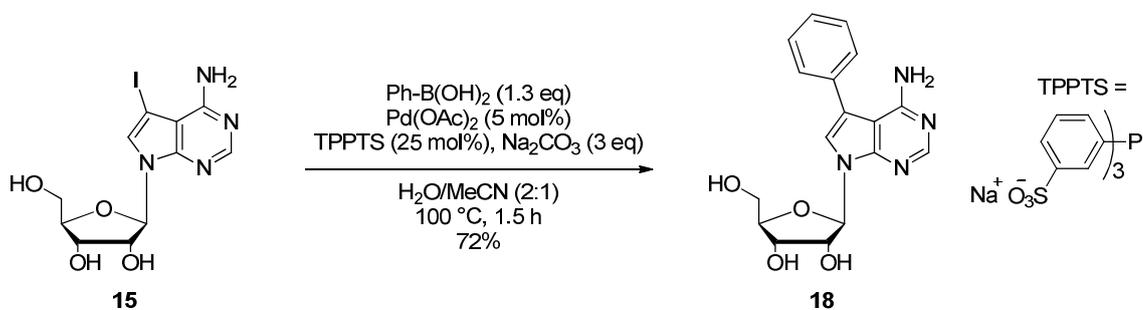
Scheme 2.10 Suzuki cross-coupling of 7-iodo-7-deazaaenosine with phenyl boronic acid.

Although these methods proved successful, the poor fluorescent properties of these analogues (see Chapter 3) emphasised the requirement for 7-deazaadenosine systems with extended π-systems.

2.3.2 Arylethynylphenyl-modified 7-deazaadenosines

Arylethynylarylene-substitued systems were identified as potentially useful "second-generation" fluorescent nucleosides. The extended π-system is expected to improve the fluorescence properties of the nucleoside (see Chapter 3). Two possible routes were identified for synthesising the target 7-[4-(arylethynyl)-phenyl]-7-deazaadenosine compounds (Scheme 2.11).
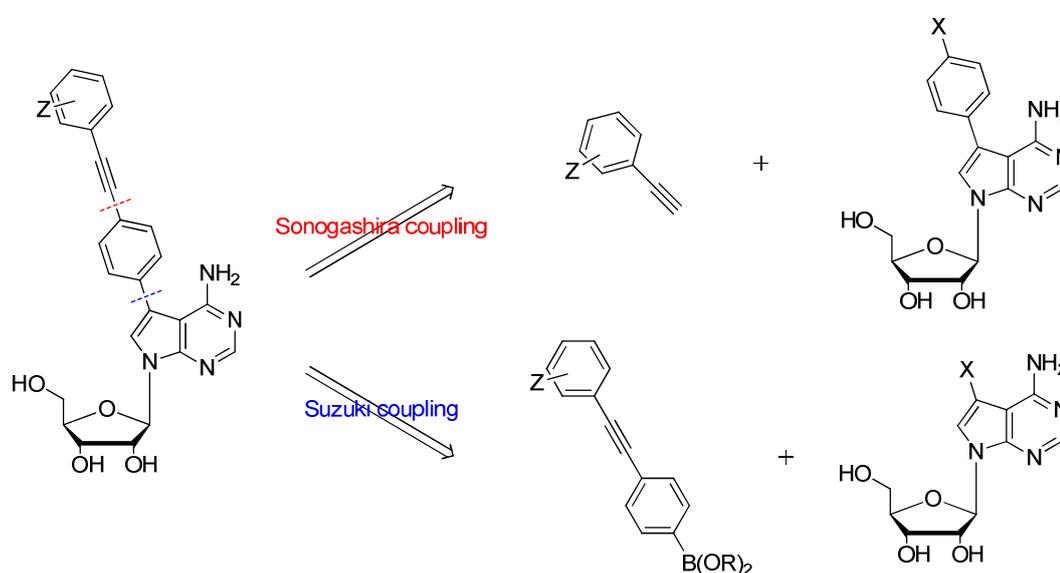
Scheme 2.11 Cross-coupling approaches to 7-arylethynylaryl-7-deazaadenosines.

7-(4-Bromophenyl)-7-deazaadenosine, **19**, was synthesised by Suzuki-Miyaura cross-coupling, in order to produce an advanced intermediate which could be further functionalised using Sonogashira cross-coupling (Scheme 2.12).
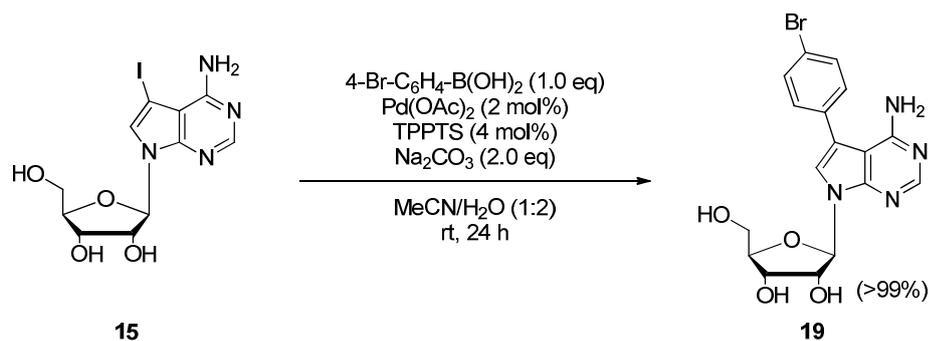
Scheme 2.12 Synthesis of 7-(4-bromophenyl)-7-deazaadenosine.

Further functionalisation of the modified nucleoside was attempted by Sonogashira cross-coupling, using the conditions developed by Firth *et al.* (Scheme 2.13).[51]
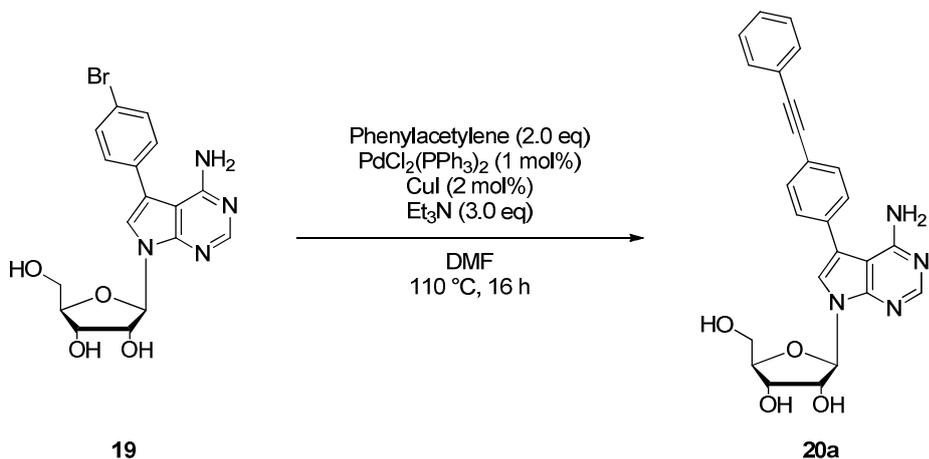
Scheme 2.13 Sonogashira coupling to synthesise target compound.

However, under these conditions the Sonogashira coupling reaction did not go to completion and the product was inseparable from the starting material by column chromatography on silica gel. The reaction was repeated using an aqueous catalyst system with tri(4,6-dimethyl-3-sulfonatophenyl)phosphine trisodium salt (TXTPS) as the ligand, and a lower temperature (80 °C), similar to the conditions developed by Shaughnessy and Western for Suzuki coupling of nucleosides,[68] and applied to the Sonogashira coupling of halogenated nucleoside triphosphates by Hocek and co-workers.[106] Subsequent to this work, Shaughnessy published a more detailed investigation of this ligand applied to the aqueous Sonogashira coupling of halogenated nucleosides, albeit with higher catalyst loadings.[107]

Under these conditions the reaction was found to be capricious, and precipitation of the catalyst as palladium black occurred on several occasions, possibly due to trace oxygen. Under strictly anaerobic conditions, the reaction proceeded to 95% completion (by NMR conversion), however as the product could not be separated from the starting material by column chromatography on silica gel this was not adequate.

Due to the difficulties encountered in the separation of this product from the 7-(4-bromophenyl)-7-deazaadenosine starting material, the second approach was attempted. This utilised Suzuki coupling to attach the arylethynylaryl fluorophore in one step. As each differently substituted boron reagent must be synthesised separately, this approach is more divergent.

Arylethynylaryl trifluoroborates have been recently synthesised by Ham and Kim, using either Sonogashira or Heck alkynylations.[108] The same conditions were applied to the synthesis of a series of arylethynylaryl trifluoroborates, **22a-e**, from the corresponding alkynes and 4-bromophenyl potassium trifluoroborate, **21** (Scheme 2.14). **21** was synthesised from the corresponding boronic acid using KHF_2, according to the procedure described by Molander and co-workers.[109]

	Ar	Crude yield /%
22a	Ph-	>99
22b	4-MeO-C_6H_4-	>99
22c	4-MeS-C_6H_4-	>99
22d	3-Thienyl-	60
22e	4-CF_3-C_6H_4-	65
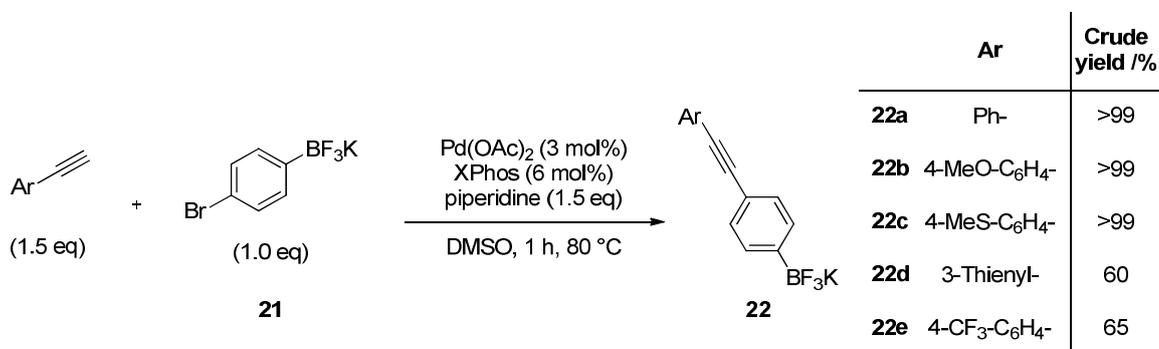
Scheme 2.14 Synthesis of Suzuki coupling partner by Heck alkynylation (yields calculated from ^{1}H NMR spectroscopic analysis).

The arylethynylaryl trifluoroborates were coupled with 7-iodo-7-deazaadenosine using Pd-catalysed Suzuki-Miyaura coupling, similar to the Shaughnessy conditions (Scheme 2.15).[68]
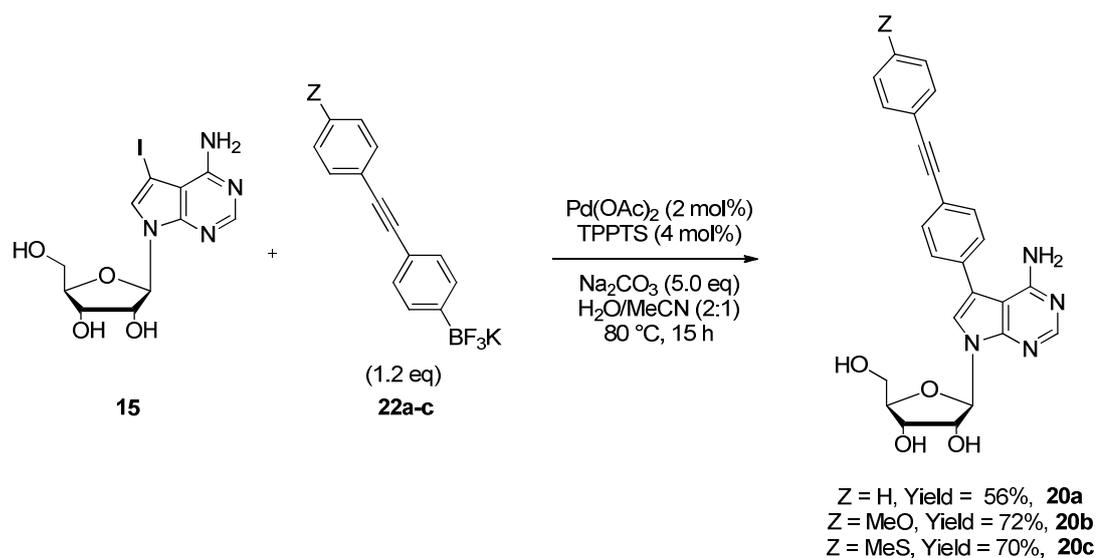
Scheme 2.15 Synthesis of target compounds by Suzuki coupling.

Although separation of the product from the starting material and organoboronic acid formed in the reaction was not trivial, the products could be isolated in moderate to good yields. This approach was used to successfully synthesise the parent (Z = H), 4-methoxy and 4-methylsulfanyl analogues. However, there are several limitations with this methodology. The major limitation is in purifying the synthesised trifluoroborate salts if the alkynylation reaction does not go to complete conversion, due to the high polarity of organotrifluoroborates and their instability to silica gel.[110] This was the case with the synthesis of the 3-thienyl and 4-trifluoromethylphenyl- substituted 4-ethynylphenyl trifluoroborates, which gave yields of 65% and 60% respectively. It was possible to couple the impure trifluoroborate products with the 7-iodo-7-deazaadenosine to give mixtures of **20** and the desired product. The mixed products could then be separated by preparative thin layer chromatography, however obviously this is far from ideal.

Due to the difficulties encountered when expanding the range of arylated nucleosides, an alternative approach was attempted. Another useful way to protect organoboronic acids is to use organoboronate esters. Following procedures published by Zheng *et al.*,[111] 4-bromophenyl boronic acid neopentylglycol ester was synthesised (Scheme 2.16) and the bromide chemoselectively coupled to phenylacetylene to give **27a** (Scheme 2.17).
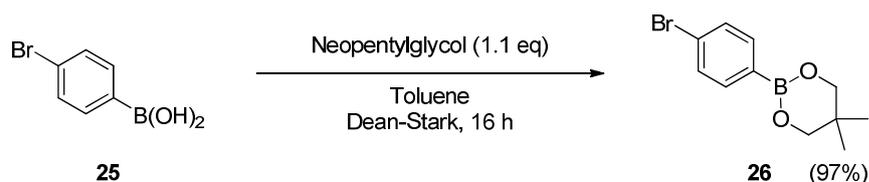
Scheme 2.16 Synthesis of neopentylglycol-protected 4-bromophenylboronic acid.
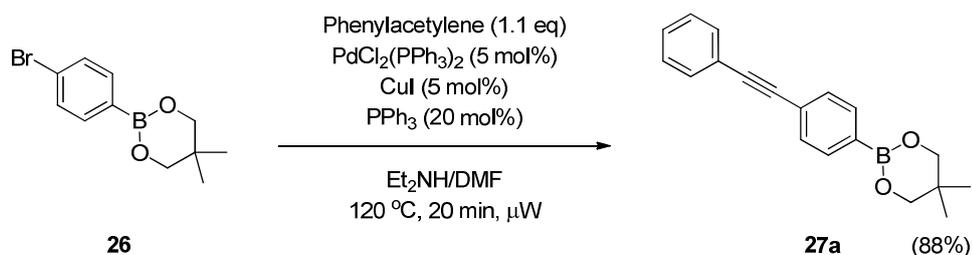
Scheme 2.17 Sonogashira coupling of 4-bromophenyl boronic acid neopentyl ester with phenylacetylene.

The bromide undergoes a Sonogashira coupling with the arylacetylene, with no observed Suzuki by-products (arising from activation of the boronate ester group). This is probably because the use of an anhydrous solvent with an amine base favour Sonogashira coupling. Although no water or aqueous base was added, and dry DMF used, no particularly rigorous air/moisture free conditions were employed in this reaction. It is thought that an aqueous base is required to activate organoboronic acids,[112] although there are some examples of Suzuki-Miyaura couplings under apparently anhydrous conditions.[113]

To develop a slightly less divergent route, 4-ethynylphenyl boronic acid neopentyl glycol ester **29** was synthesised using similar conditions. This can then be coupled to aryl halides to give the same arylethynylphenyl- products without the need to individually synthesise each arylacetylene (Scheme 2.18). The arylethynylaryl boronate esters were synthesised using one of these two routes in moderate to excellent yields (Scheme 2.19 and Table 2.2). There does not appear to be a pattern in the variation in yields, however the lower yields (Table 2.2, entries 2 and 3) may be due to these reactions being carried out on a smaller scale.
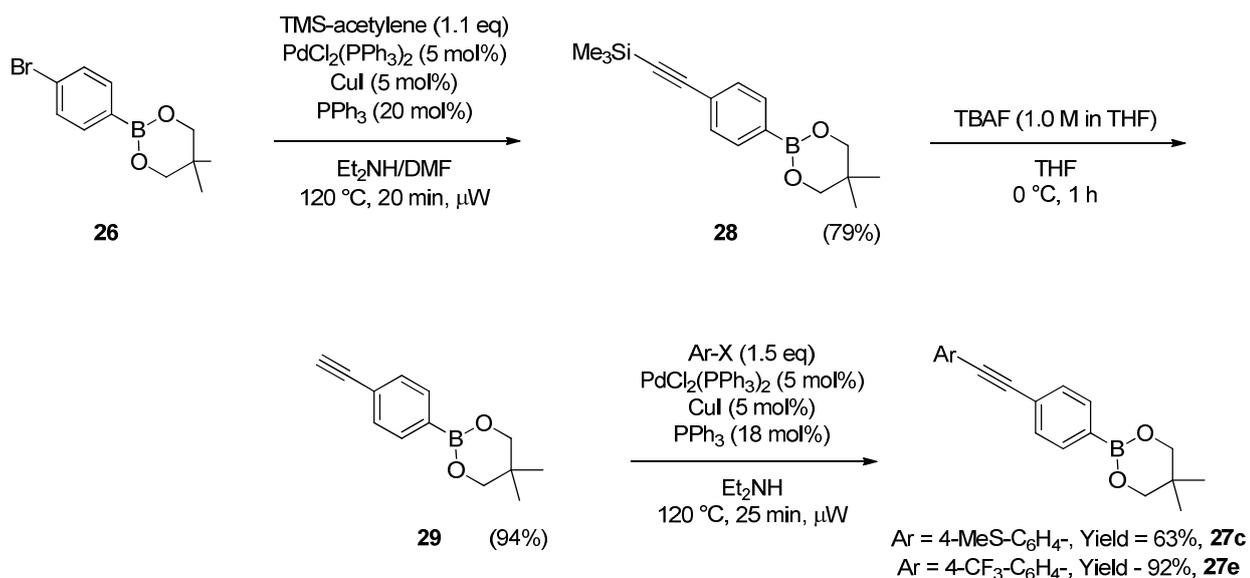
Scheme 2.18 Synthesis of phenylethynylphenyl boronate ester.

This slightly altered strategy allows for the synthesis of a wider range of fluorescent nucleosides, through greater variation of the aryl substituents. Changing the base from potassium phosphate to caesium carbonate and adjusting the metal to ligand ratio from 1:5 to 1:2.5 led to a significant improvement in the yield. A number of analogues have been synthesised in this way (Scheme 2.20 and Table 2.2).
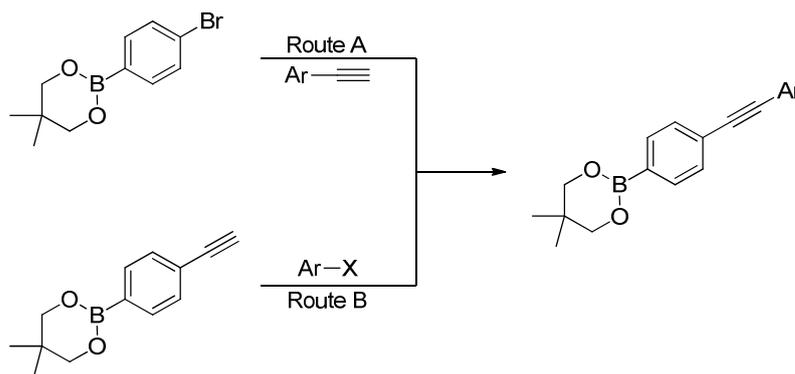
Scheme 2.19 Synthesis of arylethynylaryl nucleoside analogues.
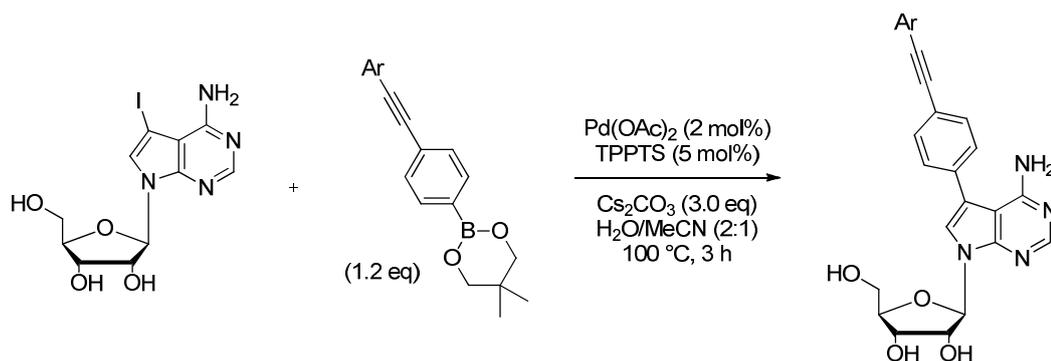
Scheme 2.20 Cross-coupling of 7-modified-7-deazaadenosines with organoboronate esters.

Table 2.2 Yields of synthesised modified nucleosides.

	Ar	Route	Yield of boronate ester[a] /%	Nucleoside Product	Yield[a] /%
1	Ph-	A	88	**20a**	74
2	4-MeO-C$_6$H$_4$-	A	62	**20b**	>99
3	4-MeS-C$_6$H$_4$-	B	63	**20c**	71
4	3-Thienyl-	A	68	**20d**	77
5	4-CF$_3$-C$_6$H$_4$-	B	92	**20e**	68

[a] Isolated yield after column chromatography on silica gel.

2.3.3 Synthesis of 2'-deoxyadenosine analogues

7-Iodo-7-deaza-2'-deoxyadenosine was synthesised according to published procedures (Scheme 2.21).[114,115] Hoffer's chlorosugar (**6**) was synthesised exclusively as the α-anomer from 2-deoxyribose in three steps (42%). It was then coupled with 6-chloro-7-iodo-7-deazapurine in a stereospecific-anion-glycosylation employing the solid-liquid phase-transfer agent TDA-1 (tris(dioxa-3,6-heptyl)amine).[116] As with the ribose analogue, a one-pot deprotection/nucleophilic aromatic substitution with aqueous ammonia and dioxane gives the product in high yield after column chromatography on silica gel.
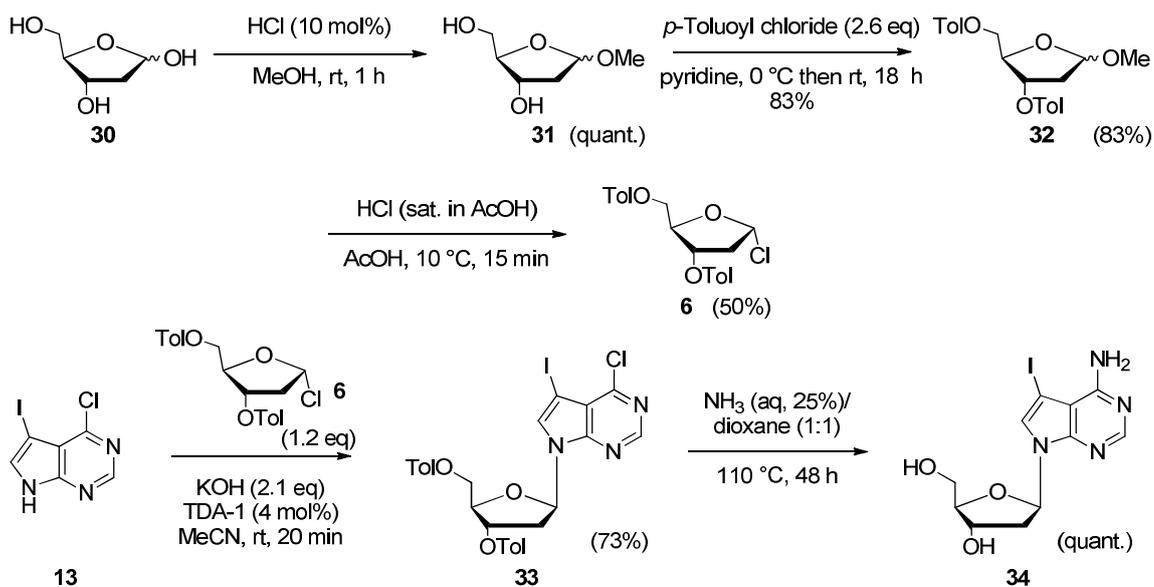
Scheme 2.21 Synthesis of 7-iodo-7-deaza-2'-deoxyadenosine.

7-Iodo-7-deaza-2'-deoxyadenosine, **34**, was then coupled with **27a** using Suzuki-Miyaura cross-coupling conditions as described above to give **35** quantitatively (Scheme 2.22).
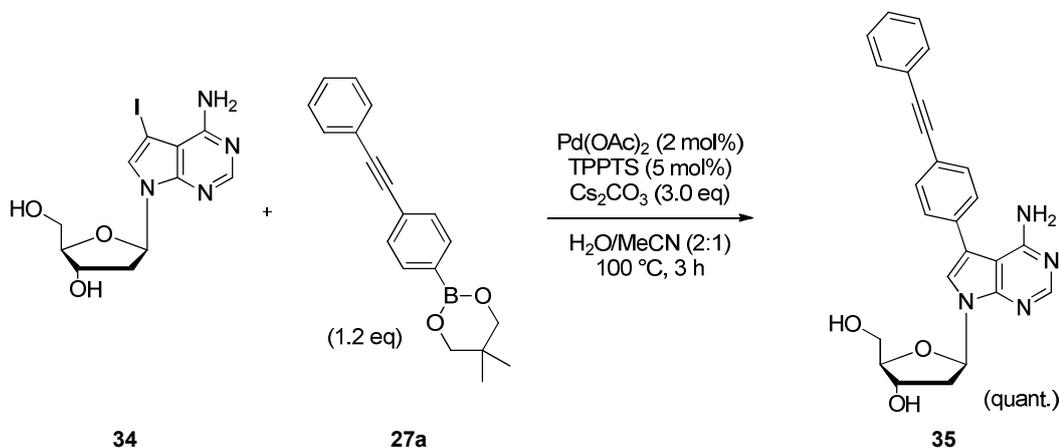
Scheme 2.22 Synthesis of the 2'-deoxyribose analogue.

The fluorescent nucleosides were comprehensively characterised using standard techniques, including NMR spectroscopy and ESI mass spectrometry, as well as photophysically by UV/vis and fluorescence spectroscopy (See Chapter 3).

2.3.4 NMR spectroscopic analysis of nucleoside products, including conformational aspects

NMR spectra of all the novel nucleosides showed a single diastereomer, consistent with the β-anomer. NMR spectra were recorded in CD_3OD or $(CD_3)_2SO$, depending on compound solubility. In particular, **20c** was found to be only partially soluble in methanol. A good ^{13}C spectrum could not be obtained for **20c** in DMSO-d_6, so HSQC was used to determine ^{13}C NMR chemical shifts. The ribose compounds were expected to display coupling between all adjacent protons on the sugar ring (additional coupling to the hydroxyls is also observed in $(CD_3)_2SO$). This would give rise to C1'-H (d, J = 6.4 Hz), C2'-H (dd, J = 6.4, 5.3 Hz), C3'-H (dd, J = 5.3, 2.8 Hz), C4'-H (ddd, J = 2.8, 2.8, 2.8 Hz), C5'-H (dd, J = 12.4, 2.8, 2.8 Hz), C5''-H (dd, J = 12.4, 2.8, 2.8 Hz). In all cases, the C4'-H was observed as an 'apparent' quartet (Figure 2.2). This would be expected if the three coupling constants (to C3'-H, C5'-H and C5''-H) were identical, however it can be deduced from the other signals that this is not the case. Instead, the observed quartet has a coupling constant equivalent to the average of the expected coupling constants. Use of a high field spectrometer did not improve the resolution of these coupling constants. The averaged signal may be the result of conformational changes occurring at room temperature. In some cases, the C2'-H is similarly observed as an apparent triplet, with an averaged coupling constant.
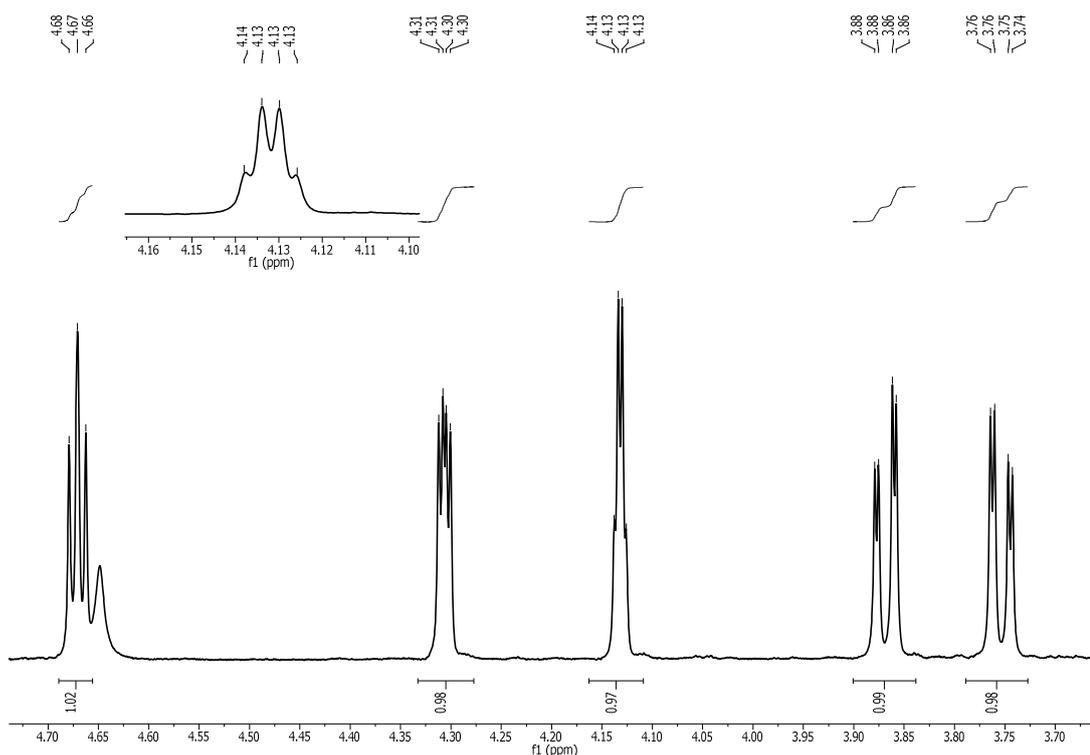
Figure 2.2 Partial ¹H NMR (700 MHz, CD₃OD) spectrum of **20a** at 298 K with inset expansion of the C4′-H signal at δ 4.13.

NMR spectroscopy is a useful tool for determining structural information about nucleosides.[117] As described previously, the ribose (or 2′-deoxyribose) sugar can adopt a number of conformations, and these can be an indicator of how similar the modified nucleoside will be to the natural nucleoside in biological systems. Sugar conformation contributes to the secondary structure of nucleic acids, so the similarity of the modified nucleoside conformation to the natural nucleoside will give an indication of any distortion of nucleic acid secondary structure caused by the fluorescent label.

The torsion angle around the nucleoside glycosyl bond χ(O4′-C1′-N9-C4) gives rise to the *syn* and *anti* conformers. The *anti* conformer (90° < χ < 270°) is the preferred conformer of the natural nucleosides at room temperature, and is the conformer that occurs in both A-type and B-type nucleic acid secondary structures. However, as discussed previously, modifications to the nucleoside can cause a switch in the preferred conformer from *anti* to *syn*.[86] Whether a compound is *anti* or *syn* can be deduced from ¹H and ¹³C NMR spectroscopy by examining the sugar chemical shifts. Generally, a switch from *anti* to *syn* is associated with an upfield shift of δ(C2′) and a downfield shift of δ(C1′), δ(C3′) and δ(C4′). The chemical shift of C2′-H can also be used to determine the *syn/anti* ratio.[118] Although

54

this method does not take into account the impact of other conformational changes on δ(C2′-H), it is still an indication of the predominant species. Examination of the spectroscopic data for the novel 7-deazaadenosine nucleosides, and comparison with the natural nucleosides, the unmodified 7-deazaadenosines and some 8-modified adenosines suggests that the compounds **20a** and **35** both preferentially adopt the *anti* conformation at 298K in $(CD_3)_2SO$ (Table 2.3).

Table 2.3 Comparison of NMR chemical shifts of natural and modified nucleosides, including the novel fluorescent 7-deazaadenosine nucleosides in DMSO-d_6.

Compound	Solvent	δ					Major
		C1′	C2′	C3′	C4′	C2′-H	conformer
A[119]	DMSO-d_6	86.7	74.3	71.4	88.8	4.64	*anti*
dA[120]	DMSO-d_6	83.8	39.7	71.8	88.7	2.73	*anti*
8-Ph-A[83]	DMSO-d_6	86.7	71.2	71.1	89.0	5.18	*syn*
8-Ph-dA[86]	DMSO-d_6	85.7	37.2	71.4	88.4	3.30	*syn*
c^{7}A[121]	DMSO-d_6	87.6	73.7	70.8	85.1	4.42	*anti*
c^{7}dA[122]	DMSO-d_6	85.8	40.2	70.8	87.8		*anti*
20a	DMSO-d_6	89.7	74.3	71.2	85.9	4.71	*anti*
	MeOH-d_4	89.9	75.6	72.5	87.3	4.67	*anti*
35	MeOH-d_4	86.5	41.5	73.0	89.1	2.72	*anti*

The second important conformational variation is the pucker of the ribose ring. There are twenty possible sugar conformations which can be described by a pseudorotational cycle (Figure 2.3).[117,123] The conformers can be divided into S-type or N-type structures. The most commonly observed and biologically important conformers are C2′-*endo* (^2E) and C3′-*endo* (^3E) conformations. C2′-*endo* conformations are associated with B-type DNA, whereas C3′-*endo* conformations are associated with A-type nucleic acid structures. By applying the Karplus equation, the ratio of C2′-*endo*:C3′-*endo* can be determined from the ^{1}H coupling constants of the nucleosides (Equation 2.1).[123] This method assumes that only C2′-*endo* and C3′-*endo* conformations are present. The coupling constants relevant to the sugar pucker are $J_{3'4'}$ and $J_{1'2'}$. As the C4′-H signal does not have resolved coupling constants (see above), $J_{3'4'}$ was measured from the C3′-H signal. The novel fluorescent 7-deazaadenosines have C2′-*endo*:C3′-*endo* ratios close to those observed for the natural adenosines and unmodified 7-deazaadenosines. This is a good indicator that the modifications introduced

will not significantly disrupt the secondary structure of duplexes containing these nucleosides or change their biological properties through changes in sugar conformation.

$$\%C2'\text{-}endo = 100(\frac{J_{1'2'}}{J_{1'2'} + J_{3'4'}})$$

Equation 2.1 Determination of the percentage of C2'-*endo* conformation.
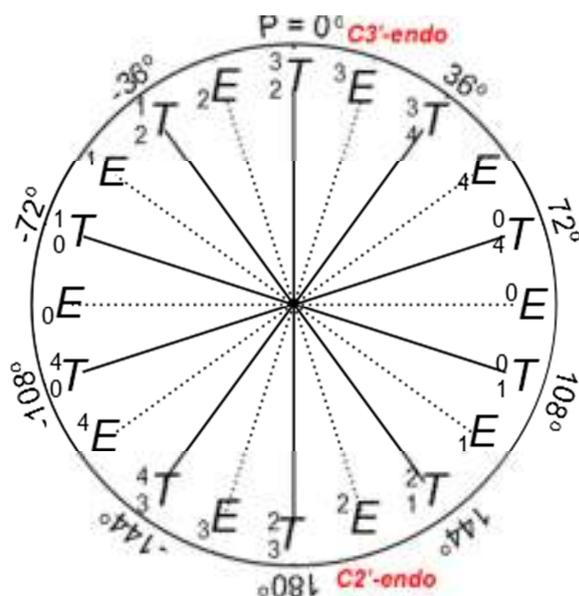
Figure 2.3 Pseudorotational cycle highlighting the commonly observed sugar puckers.

Table 2.4 Coupling constants and ratio of C2'-*endo* to C3'-*endo* for novel 7-deazaadenosines and comparison with other adenosine compounds.

Compound	Solvent	$J_{1'2'}$	$J_{3'4'}$	%C2'-*endo*	%C3'-*endo*
A	DMSO-d_6	6.2	3.0	67	33
dA	DMSO-d_6	7.9	2.6	75	25
7-I-c[7]A	DMSO-d_6	6.4	m	n.d.	n.d
7-I-c[7]dA	DMSO-d_6	8.2	2.6	76	24
20a	DMSO-d_6	6.2	m	n.d.	n.d.
	MeOH-d_4	6.4	2.8	70	30
35	MeOH-d_4	7.1	2.7	72	28

2.4   Phosphoramidite synthesis

2.4.1   Introduction

A common method for incorporating unnatural nucleosides into oligonucleotides is by chemical solid-phase synthesis.[124,125] This has the advantage of tolerating many highly modified nucelobases and sugars which would not normally be tolerated by nucleic acid polymerase enzymes. [126]   The presence of only one secondary hydroxyl in deoxyribonucleosides makes the synthesis of the 3'-phosphoramidite significantly more straightforward for DNA than RNA. [127]   Modern fully-automated systems mean oligonucleotides can be synthesised with very little error.

A   3'-phosphoramidite   (generally   derived   from   2-cyanoethyl-*N,N*-diisoproylchlorophosphoramidite) of the nucleoside is synthesised which acts as a precursor to the phosphate group. Appropriate protecting groups are also required for each nucleoside; usually the 5'-hydroxyl is protected with a 4,4'-dimethoxytrityl group which can be removed orthogonally from common amino and hydroxyl protecting groups such as benzoyl and silyl groups; the method is selective for primary hydroxyls.[125]

Nucleic acids are synthesised from the 3'-end to the 5'-end. First, the 3'-hydroxyl of the first nucleoside is attached to a solid-support (usually through a non-nucleoside linker such as succinyl, attached to the 3'-O position of the first nucleoside). The trityl protecting group on the 5'-position is then removed under acidic conditions (Scheme 2.23, **I**). The free hydroxyl can then couple to another nucleoside phosphoramidite (**II**). Any unreacted 5'-hydroxyls can be capped using acetic acid and 1-methylimidazole to prevent reaction in the next cycle of the synthesis (**III**), and to allow easy removal of these shorter oligonucleotides at the end of the synthesis. The P^{III} phosphoramidite is then oxidised to the P^{V} phosphate that makes up the backbone of nucleic acids (**IV**).

The cycle of detritylation, coupling, capping and oxidation is repeated until all the desired nucleobases have been incorporated. The oligonucleotide can then be cleaved from the solid-support.
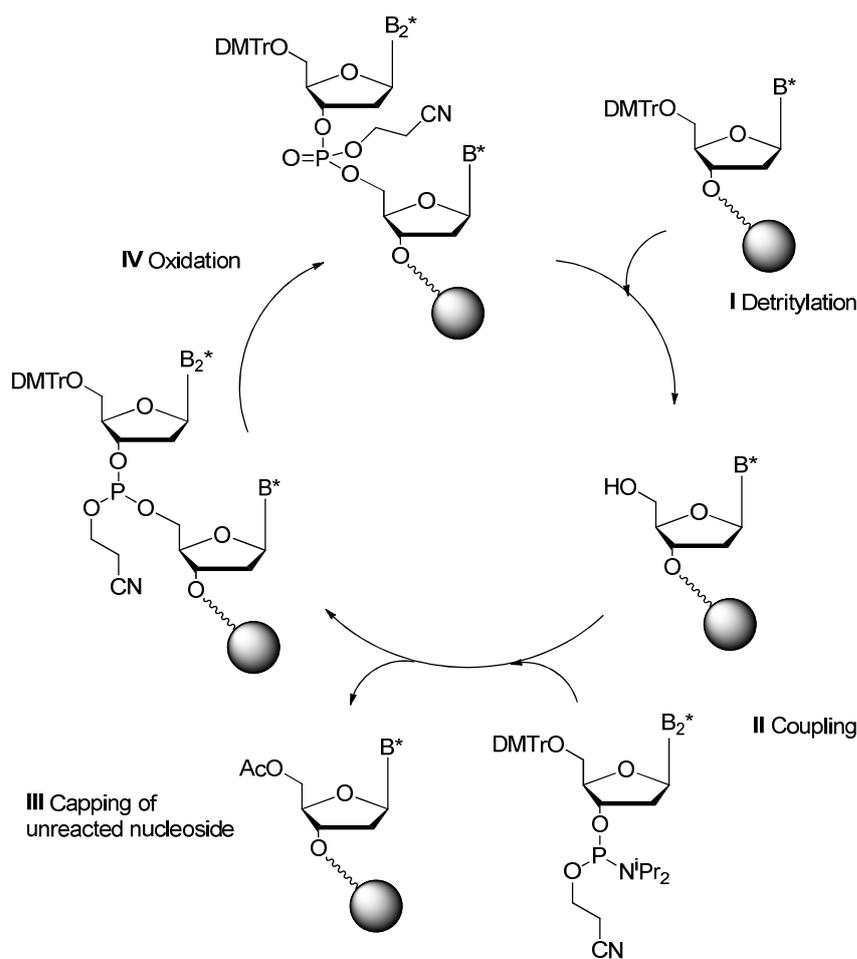
Scheme 2.23 Solid-phase synthesis of oligonucleotides.

2.4.2 Synthesis of fluorescent nucleoside phosphoramidite monomer

The parent compound, containing the 4-(phenylethynyl)phenyl group at C7 (**35**), was chosen for incorporation into oligonucleotides due to its simple structure and promising fluorescent properties. The most commonly used protecting group for the exocyclic amine of adenosine nucleosides prior to solid phase synthesis is benzoyl, however other groups are also sometimes used. For incorporation of the parent compound into oligonucleotides using solid-phase synthesis, the amine was protected as a dimethyl formamidine by reaction with dimethylformamide dimethyl acetal (DMF-DMA). This protecting group was chosen as it has been used previously with 7-modified-7-deazaadenosines;[128,129] recently it has been shown to impart greater stability to unnatural adenine nucleoside phosphoramidites.[130] It is also easy to install, does not require prior protection of the sugar hydroxyls, and can be removed under standard oligonucleotide deprotection conditions (*i.e.* aqueous ammonia). The formamidine was formed quantitatively, which allowed the

crude product (after removal of the solvent) to be used without further purification. The second protection step, installation of the 5'-hydroxyl with DMTrCl, was also successful, giving **36** in 73% yield over two steps (Scheme 2.24).
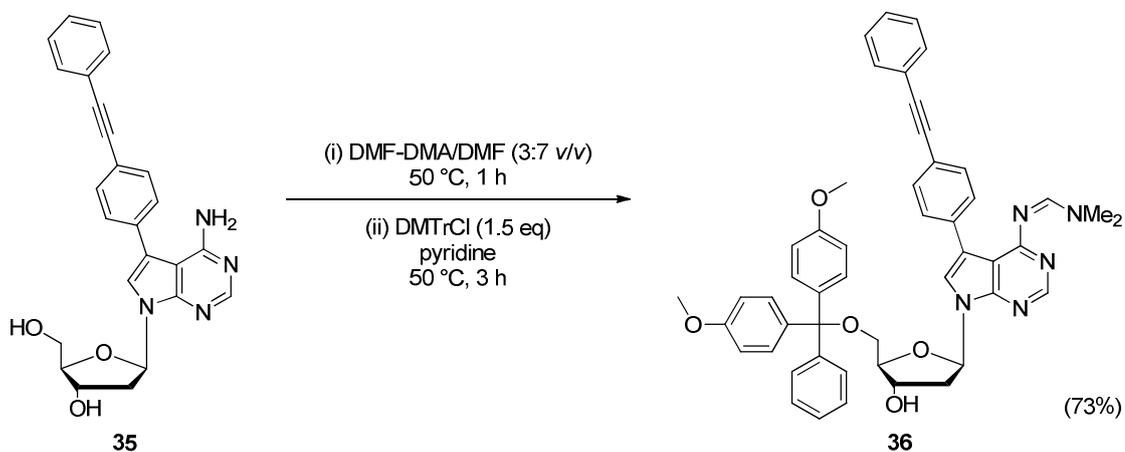
Scheme 2.24 Protection of **35** for solid-phase synthesis.

The protected nucleoside was then converted to the phosphoramidite, **37**, by reaction with 2-cyanoethyl-*N*,*N*-diisoproylchlorophosphoramidite to give a mixture of diastereomers (Scheme 2.25). Unusually for nucleoside phosphoramidites of this type, the diastereomers have similar resonances in the ^{31}P NMR spectrum and are thus virtually indistinguishable (Figure 2.4).   An unknown impurity which was present in the 2-cyanoethyl-*N*,*N*-diisoproylchlorophosphoramidite starting material (^{31}P NMR δ 14.5) was also present in the product, but did not have any impact on the solid-phase synthesis.
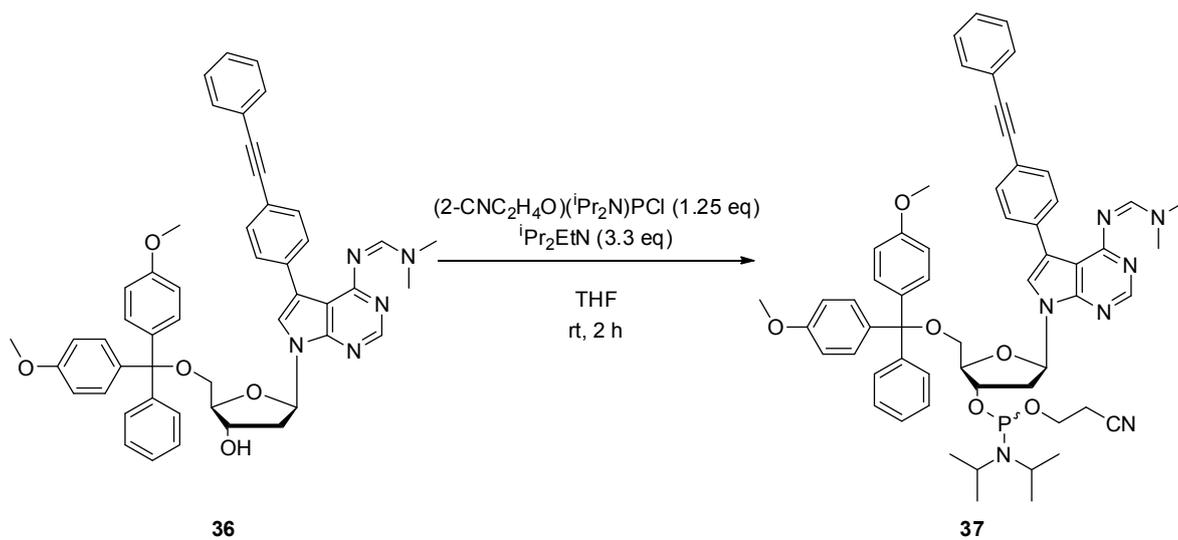
Scheme 2.25 Synthesis of nucleoside phosphoramidite **37**.
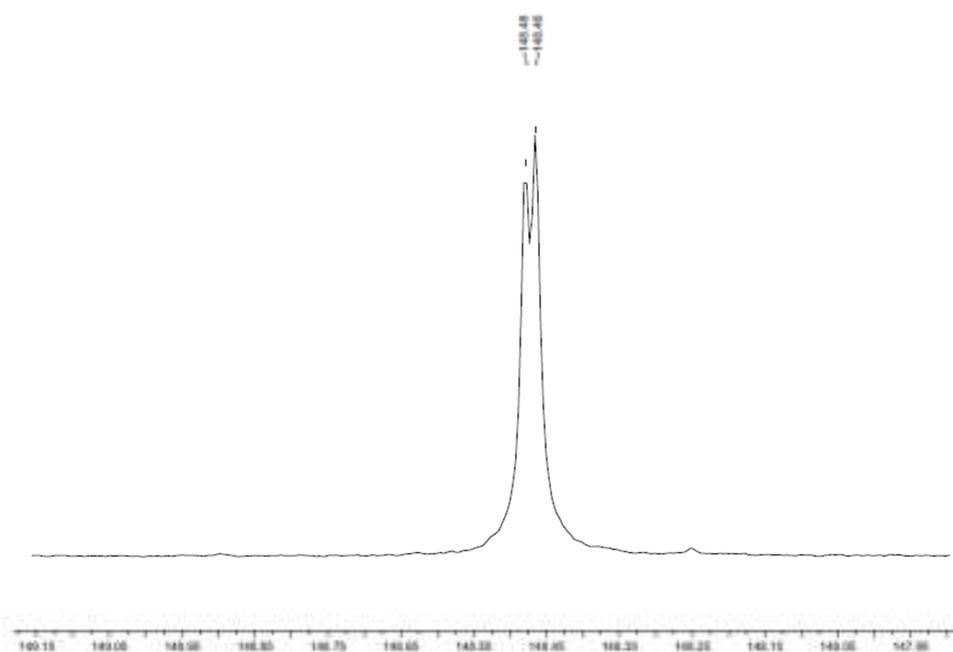
Figure 2.4 ^{31}P NMR (212 MHz, CD$_2$Cl$_2$) spectra of **37**.

The synthesised phosphoramidite was incorporated into a single position of five different 13-base oligomers (**ODN1→ODN5**) by BioTeZ Berlin-Buch GmbH. The purity of the synthesised oligonucleotides was gauged by positive-mode MALDI-TOF-MS (Table 2.5).

Table 2.5 Synthesised oligonucleotides calculated and found MALDI-TOF-MS.

	Sequence[a]	Calculated MW [M+H]	Found MW (MALDI)[b]	Error /ppm	% GC content
ODN1	$^{5'}$d(CGCTTG**a**GTTCGC)$^{3'}$	4115.76	4115.8227	16.15	62
ODN2	$^{5'}$d(CGCTTC**a**CTTCGC)$^{3'}$	4035.74	4035.8191	18.62	62
ODN3	$^{5'}$d(CGCAAG**a**GAACGC)$^{3'}$	4151.80	4151.8688	15.98	62
ODN4	$^{5'}$d(CGCAAC**a**CAACGC)$^{3'}$	4071.79	4071.8898	24.46	62
ODN5	$^{5'}$d(CGCAAT**a**TAACGC)$^{3'}$	4101.79	4101.8589	16.92	46

[a] **a** = **35** residue [b] 3-Hydroxypicolinic acid (3-HPA) in 1:1 MeCN:H_2O (0.5 M) was used as the matrix and mixed 1:1 with the oligonucleotide dissolved in water (*ca.* 500 μM)

These sequences were selected to investigate the effect of the local environment on the modified base by varying the neighbouring bases. Purine-rich (**ODN3**) and pyrimidine-rich (**ODN1** and **ODN2**) sequences, and the effect of GC rich (**ODN1→4**) and AT rich (**ODN5**) oligomers were chosen. The properties of the oligonucleotides are discussed in Chapter 4.

2.5   Phosphorylation

2.5.1   Introduction

Triphosphorylation of modified nucleosides is required prior to incorporation into nucleic acids by polymerase enzymes. Nucleotide triphosphates are extremely unstable, even at low temperatures, which makes their synthesis, storage and handling problematic.[131] Although much research has been conducted into the synthesis of nucleotide triphosphates, from the corresponding nucleosides, no general synthetic method has yet been developed.

Routes to nucleotide triphosphates usually require reaction of the nucleoside with an electrophilic phosphorous centre. If this is a monophosphate precursor (such as $POCl_3$), it is then reacted further with a nucleophilic pyrophosphate. One of the most widely applicable routes to nucleotide triphosphates was developed independently by Ludwig[132] and Ruth and Cheng.[133] Building on work by Yoshikawa and co-workers,[134] they developed a one-pot synthesis of nucleotide triphosphates from the corresponding unprotected nucleosides

(Scheme 2.26). This methodology has been found to be one of the most widely applicable phosphorylation strategies, particularly for the synthesis of unnatural nucleosides.[135]
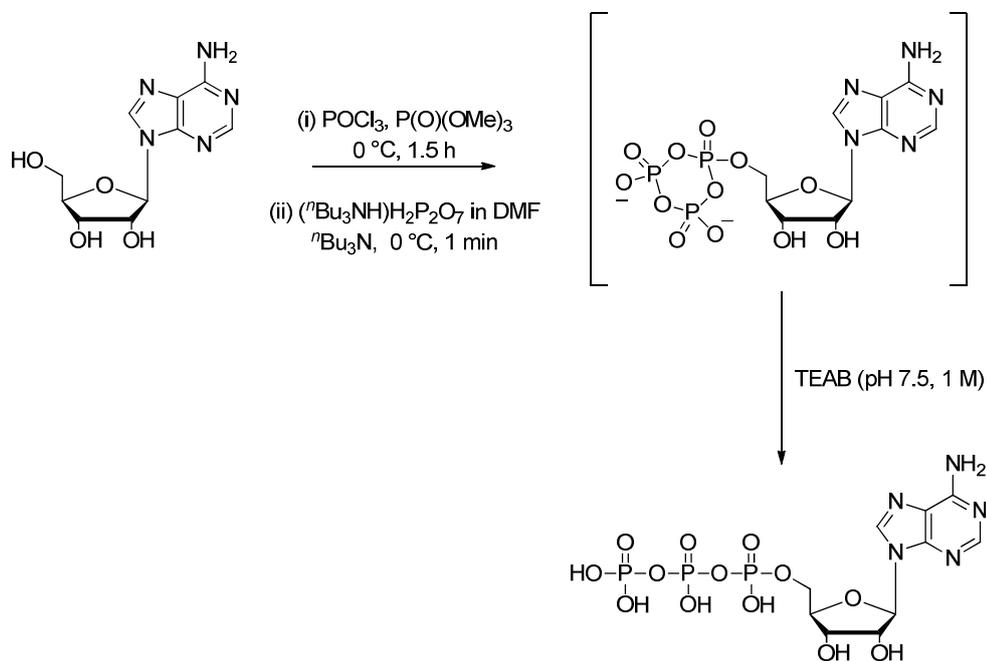
Scheme 2.26 Ludwig's one-pot synthesis of ATP from unprotected adenosine.

It was found that the use of trialkylphosphate solvents was key to the selectivity for 5'-triphosphates, suggested to be due to the formation of a complex between the solvent and the nucleoside which activates the 5'-position (Figure 2.5).[136] However, recently an example of a phosphorylation procedure which gives high selectivity for the 5'-position in a more conventional organic solvent, acetonitrile, has been published.[137]
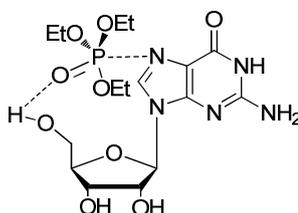
Figure 2.5 Proposed complex formed between triethylphosphate and guanosine.

Another problem with the POCl$_3$-based syntheses is the formation of the Vilsmeier reagent (Scheme 2.27, **I**) from POCl$_3$ and DMF.[138] This activated reagent can go on to form 2',3'-

bicyclic phosphate nucleotides (**III**) or inorganic triphosphate (**II**). 2′,3′-Bicyclic phosphates can also be formed from the decomposition of 2′- or 3′-triphosphates.[132]
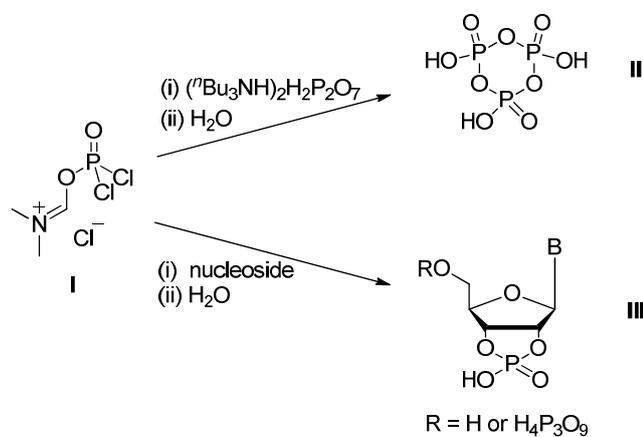
Scheme 2.27 Byproducts formed during phosphorylation.

A common alternative route to nucleotide triphosphates was developed by Ludwig and Eckstein. [139] This approach uses salicyl chlorophosphite (2-chloro-1,3,2-benzodioxaphosphorin-4-one, **39**) as the α-phosphorus component, which is coupled to the nucleoside before addition of pyrophosphate and oxidation of the α-phosphorus to P^V (Scheme 2.28).
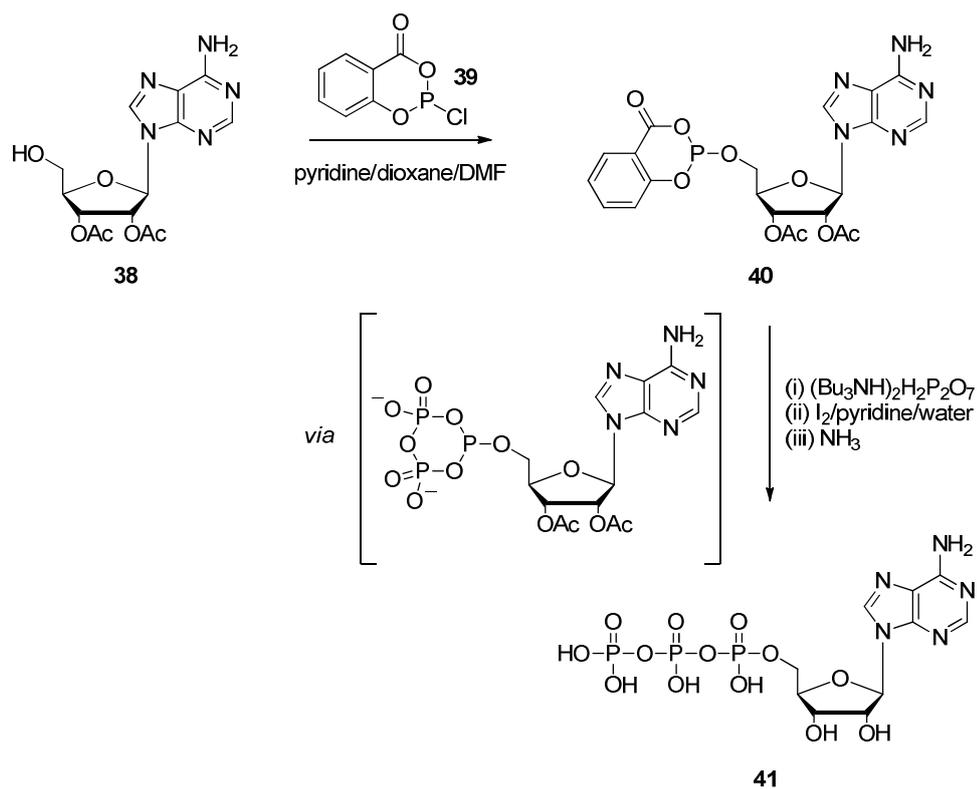
Scheme 2.28 Ludwig-Eckstein synthesis of ATP from salicyl chlorophosphite.

This route is advantageous because it allows synthesis of the α-thiophosphate (NTPαS) by oxidation with sulfur instead of iodine/water. It also gives very good selectivity for the triphosphate product over mono- or diphosphates, however protection of the 2′- and 3′-hydroxyls is usually necessary as selectivity for the primary hydroxyl over the secondary hydroxyls is poor. Recently, this synthesis has been improved upon by Huang and co-workers for the synthesis of ribonucleotides[140] and deoxyribonucleotide.[141] Their modification improves the selectivity for triphosphates over mono- and diphosphates by pre-forming the triphosphate reagent (**42**) by reaction of salicyl chlorophosphite with pyrophosphate prior to introduction to the nucleoside (Scheme 2.29). Another benefit is the use of unprotected nucleosides in this chemistry. The ratio of 5′-phosphorylated products to 2′- and 3′-triphosphorylated products was found to be *ca.* 85:15 for ATP and *ca.* 90:10 for dATP (although the crude yield of ATP was only 10%). The nucleotide triphosphates can be used crude or further purified by HPLC, but unlike the POCl₃-based syntheses the products do not require an FPLC purification step.
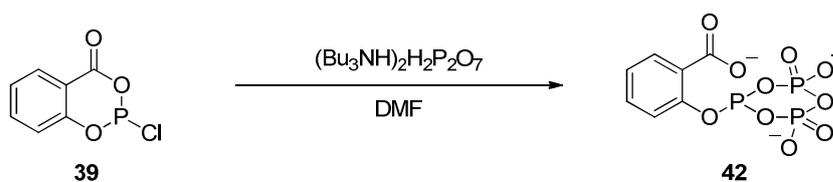
Scheme 2.29 Synthesis of triphosphitylating reagent from salicyl chlorophosphite.

2.5.2 Phosphorylation using Ludwig conditions

The phosphorylation of the parent compound **20a** was attempted using both methods described above. The phosphorylation was carried out using a procedure developed in-house based on the Ludwig conditions (Scheme 2.30).[142]
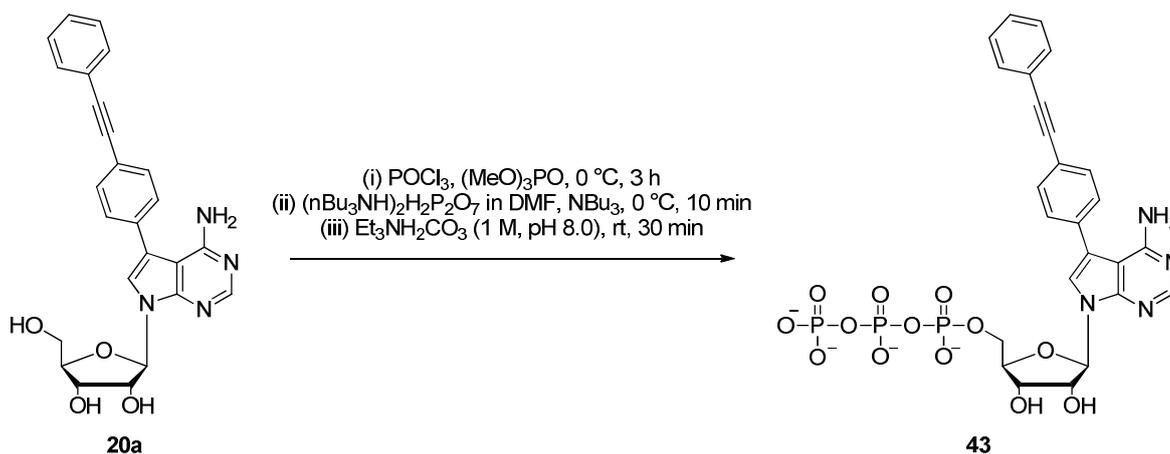
Scheme 2.30 Attempted synthesis of nucleotide triphosphate using Ludwig conditions.

The crude reaction mixture was dried under vacuum and then purification by FPLC (DEAE Sephadex) was attempted. Several problems were encountered at this stage of the synthesis. Elution of the triphosphate from the ion-exchange column using a gradient up to 1 M TEAB (Et$_3$NH$_2$CO$_3$ buffer, pH = 8.0) was not very successful due to the extremely long retention time and poor separation of the triphosphate from higher phosphates (Figure 2.6). 8-Phenyladenosine triphosphate elutes at *ca.* 0.6 M TEAB on a similar gradient, suggesting that the increased lipophilicity of these modified nucleosides may be the cause of the problems with the elution and separation by FPLC.[142] The buffer concentration was increased to 1.2 M but this did not lead to a significant improvement in the retention time

or peak breadth. The very broad peaks may be an indication that the modified nucleotide is too non-polar, and therefore close to precipitation.
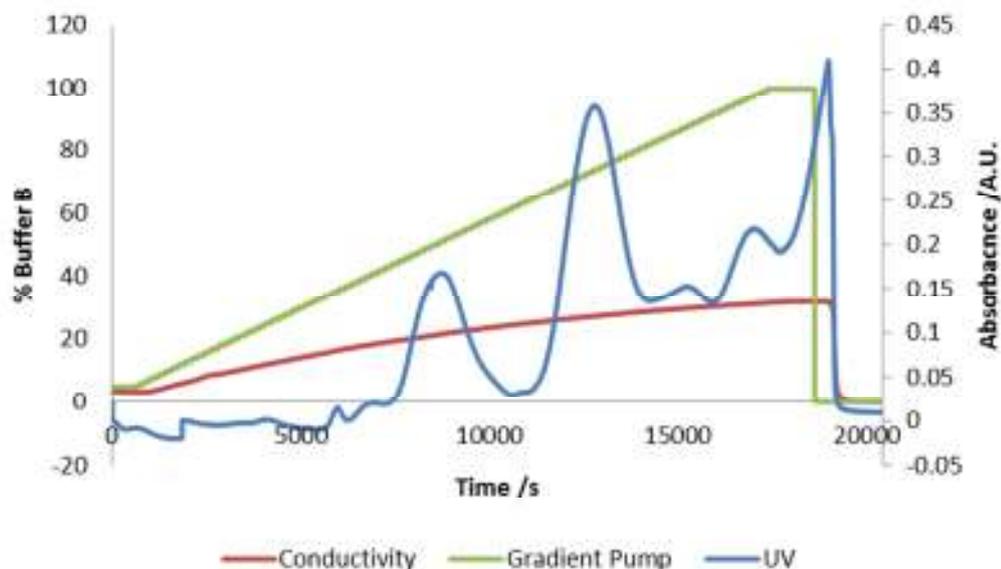
Figure 2.6 Example of an FPLC trace from the purification of NTP **43**, the final peak (still eluting at the end of the run) is thought to be the triphosphate.

Another problem encountered at this stage of the synthesis was the decomposition of the product during purification. Due to the large volumes of product-containing fractions collected after FPLC purification, lyophilisation was not a practical option for obtaining the solid product. Therefore, the product-containing fractions were dried using a rotary evaporator, keeping the water bath at <50 °C. Whilst this ensures that the triphosphate product is not completely destroyed, it does result in some decomposition. At this stage, the product could be identified by mass spectrometry. The product must then be purified further by HPLC, at which stage further decomposition can occur, and indeed no product was successfully identified after HPLC purification and solvent removal.

2.5.3   Phosphorylation using Huang conditions

Phosphorylation of the parent nucleoside compound was also attempted using Huang's conditions.[140] Although Huang and co-workers showed that selectivity for the 5'-position

was very good, even at room temperature, the selectivity using the 7-modified nucleoside was only moderate when the reaction was carried out at 0 °C (Scheme 2.31).
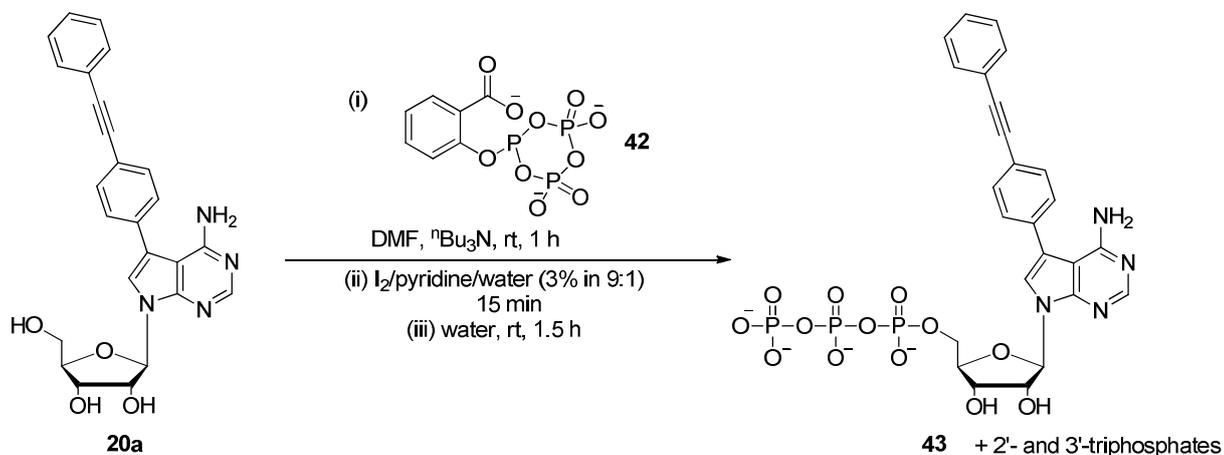
Scheme 2.31 Phosphorylation using Huang conditions.

2.6 Conclusions

A novel class of 7-deazaadenine nucleosides, both ribo- and deoxyribo-analogues, were synthesised using Pd-catalysed cross-coupling chemistry. 4-(Arylethynyl)phenyl boronate ester reagents were synthesised using chemoselective Pd-catalysed Sonogashira coupling of 4-bromophenyl boronic acid neopentyl glycol ester with terminal acetylenes. The 4-(arylethynyl)phenyl boronate esters were then coupled with 7-iodo-7-deazaadenine nucleosides under aqueous Suzuki-Miyaura conditions, using the water-soluble phosphine ligand TPPTS. Cs_2CO_3 was found to be a superior base to K_3PO_4 or Na_2CO_3, and the ligand to metal ratio was also found to be key in achieving good yields. The nucleosides were isolated in good to excellent yields (68% to quant.).

The parent deoxyribo-compound (**35**) was prepared for solid-phase oligonucleotide synthesis, with the exocyclic amine protected as a dimethyl formamidine group, and the 5′-hydroxyl protected with a 4,4′-dimethoxytrityl (DMTr) group. Phosphitylation of the protected nucleoside gave the nucleoside 3′-phosphoramidite, which was commercially incorporated into a series of oligonucleotides.

Attempts to synthesise the nucleoside 5′-triphosphate were largely unsuccessful due to the difficulties encountered in purifying the product. Using Huang and co-workers' methodology resulted in the formation of triphosphate product, but this was not successfully separated from the other nucleoside phosphate products formed.

3 Photophysical characterisation of modified nucleosides

3.1 Introduction

3.1.1 Principles of photophysics

Of key importance in the synthesis of fluorescent biological probes are their photophysical properties. To be useful in the study of biochemical and biophysical interactions, the compounds must have appropriate spectral properties. Both the UV-Vis and fluorescence properties must be evaluated to determine the effectiveness of the fluorophore. The absorption properties of the chromophore are governed by the Beer-Lambert Law (Equation 3.1). The molar absorption (or extinction) coefficient (ε) (a measure of how efficiently the compound absorbs light) can be calculated by measuring the absorbance (A) over a range of concentrations (c).

$$log \frac{I_0}{I} = A = \varepsilon c l$$

Equation 3.1 Beer-Lambert Law.

I_0 = intensity of incident light, I = intensity of transmitted light, l = path length.

The absorption wavelength must be shifted away from the protein and nucleic acid absorptions at 280 nm and 260 nm so the fluorophore can be selectively excited. The emission wavelength must then also be shifted away from the emissions of the aromatic amino acids (295-350 nm) so the fluorophore can be easily detected.

UV-Vis absorption occurs when an electron is excited from the HOMO (S_0) to the LUMO (S_1) by a photon. The absorbance wavelength of a chromophore is therefore dependent on the energy gap between the HOMO and LUMO. The higher the energy of the transition between the HOMO and the LUMO, the shorter the wavelength of light required for the transition. UV-Vis transitions can occur between any symmetry allowed combination of orbitals, for example, $\pi \rightarrow \pi^*$, $n \rightarrow \pi^*$ or $n \rightarrow \sigma^*$. Intramolecular charge transfer (ICT)

transitions can also occur, whereby the excitation induces a change in charge distribution across the molecule.[143] These usually occur when the chromophore contains electron donor and electron acceptor groups which are in conjugation with each other. The electronic transitions which can occur can be illustrated using a Jablonski diagram (Figure 3.1).[144]
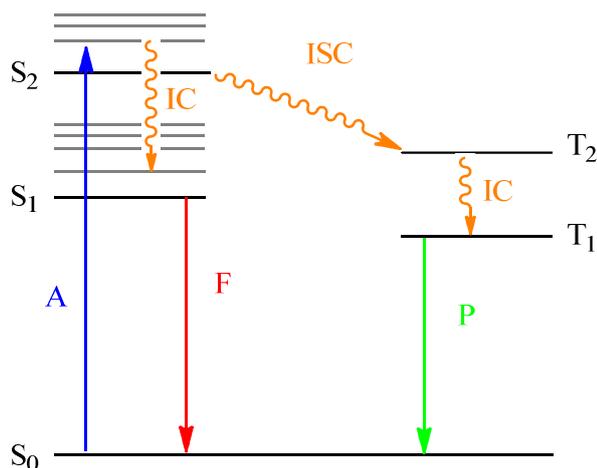
Figure 3.1 Jablonski diagram showing electronic transistions.
A = absorbance, F = fluorescence, P = phosphorescence, ISC = intersystem crossing, IC = internal conversion.

If some of the vibrational levels of the ground and excited states overlap, the excited electron can decay through non-luminescent (non-radiative) processes (internal conversion). Solvent interactions can then dissipate the energy through collisions with the chromophore, or the energy can be lost through vibrational relaxation. Luminescent emission occurs when an electron excited to S_1 relaxes by emitting a photon (Figure 3.2).
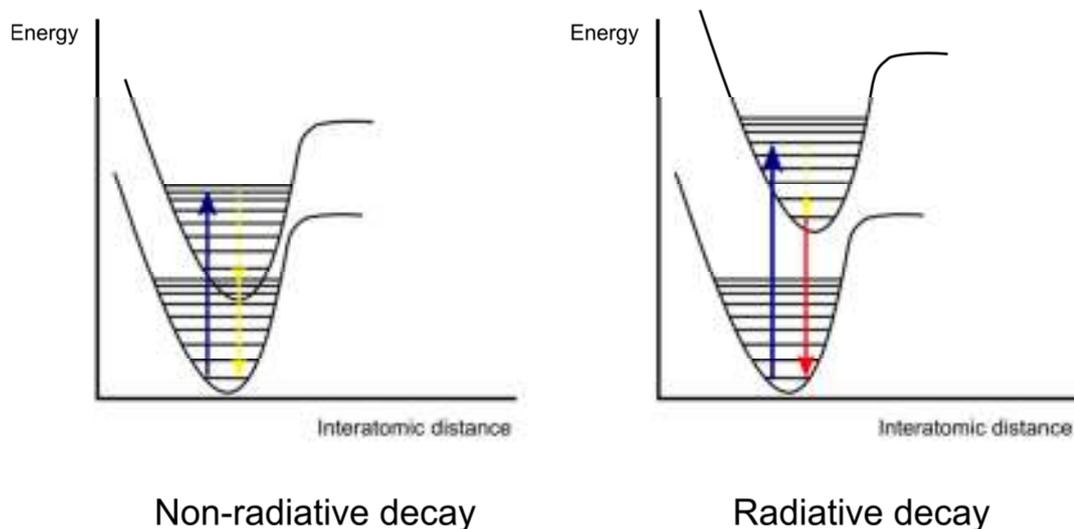
Figure 3.2 Diagram showing radiative and non-radiative decay of electronic excited states.

Luminescent emission can be either fluorescence (decay from a singlet state) or phosphorescence (decay from a triplet state, following intersystem crossing). Luminescent emission is generally bathochromically shifted with respect to the absorbance wavelength. The difference between the absorbance and emission wavelengths is called the Stokes shift, and is due to energy loss through internal conversion (usually to a lower vibrational energy level). The Stokes shift is generally measured in wavenumbers, as this is proportional to energy.

Two key properties of any fluorophore are the quantum yield and the fluorescence lifetime. Quantum yield (Φ) is a measure of the efficiency of a fluorophore, *i.e.*, it is a measure of what percentage of the photons absorbed are emitted as fluorescence. It can be expressed in terms of the emissive rate (Γ) and the rate of non-emissive decay (k_{nr}) (Equation 3.2).[145] Fluorophores with higher quantum yields are more efficient, and many commercial fluorophores have quantum yields close to 1.

$$\Phi = \frac{\Gamma}{\Gamma + k_{nr}}$$

Equation 3.2 Equation governing quantum yield.

The lifetime of a fluorophore (τ) is the average time it spends in the excited state before decay to the ground state. It can also be described as the time taken for the intensity to decay to 1/e of the initial value (I_0). The lifetime of the excited state can be used to differentiate between fluorescent and phosphorescent emission. Fluorescence lifetimes are short (*ca.* 10^{-8} s), whereas phosphorescent lifetimes are usually much longer (milliseconds to seconds) due to the forbidden transition between the triplet excited state and the singlet ground state. The fluorescent lifetime is also dependent on the rates of both emissive and non-emissive decay (Equation 3.3).

$$\tau = \frac{1}{\Gamma + k_{\mathrm{nr}}}$$

Equation 3.3 Equation governing fluorescence lifetime.

The time-dependence of fluorescence decay can be described by an exponential (or multi-exponential) decay (Equation 3.4). Fluorescence lifetimes can therefore be calculated by plotting log[I(t)] against t.

$$I(t) = I_0 e^{\frac{-t}{\tau}}$$

Equation 3.4 Function for a single exponential fluorescence decay.

I(t) = Intensity at time t.

3.1.2 Fluorescent probes

Fluorescent probes can be divided into two categories: intrinsic fluorophores and extrinsic fluorophores. Intrinsic fluorophores are those arising from naturally occurring fluorescence present in the system being studied. In biological systems, this is often due to tryptophan residues in proteins. Extrinsic fluorophores are fluorescent probes which are introduced in to a system in order to study it. Extrinsic fluorophores can be covalently or non-covalently bonded to the system they are being used in. When studying biological systems using an

extrinsic fluorophore, it is important to know whether the fluorescent label influences the system being studied, as attaching large or bulky groups may affect the biological activity.

Fluorophores generally contain extended conjugated π-systems, as the energy gap between the ground and excited states lowers as the conjugation increases. The relationship between conjugation and absorbance maxima was quantified for conjugated dienes by Woodward and Fieser in the 1940s.[146] Lowering the energy gap increases the absorption wavelength, which helps shift the absorbance away from the background amino acid and nucleotide absorbances.

Most commonly used organic fluorophores are based around aromatic cores, such as naphthalene and pyrene, or heteroaromatic cores, such as coumarin and pyrrole (Figure 3.3).
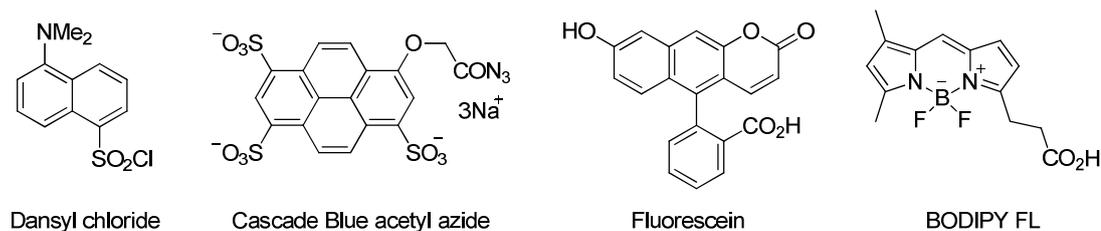
Figure 3.3 Examples of common extrinsic fluorophores.

Previous work in the group on 8-modified-adenosines has shown that analogues with directly linked aromatic groups such as 8-thienyl- and 8-phenyl-adenosines show increased fluorescence (higher quantum yields) compared with arylethynyl substituted analogues.[91] However, the analogues with ethynyl-linked aryl groups showed bathochromically shifted absorption and emission maxima compared with their directly linked counterparts, due to the extended π-system, often making them more effective for use in real systems.

3.2   UV-Vis spectroscopy

UV-Vis spectra of the 7-modified-7-deazaadenosines were recorded in DMSO over five concentrations between *ca.* $1x10^{-6}$ M and $1x10^{-4}$ M, and the Beer-Lambert law used to calculate the molar absorption coefficients (Table 3.1). DMSO was selected as a solvent for comparison to the previously synthesised 8-modified purine nucleosides, and because of the good solvation of nucleosides in DMSO. The UV-Vis spectra show, as expected, that increasing the size of the conjugated system causes a bathochromic shift in the absorption wavelength.   So 7-phenyl-7-deazaadenosine absorbs at a lower wavelength than 7-(4-[phenylethynyl]phenyl)-7-deazaadenosine (286 nm and 321 nm respectively).   The nucleosides with diphenylacetylene-based substituents (entries 4-6, 8) also exhibit a second peak (as a shoulder) in the UV spectrum which is hypsochromically shifted with respect to the main peak. The two peaks in the UV spectrum may be due to two different electronic transitions, most likely $\pi{\rightarrow}\pi^*$ and intramolecular charge transfer (ICT). The UV spectrum of the thienyl-containinly compound (entry 7) is more complex, with several other peaks present at shorter wavelengths. The same is true of the napthylethynyl-substituted nucleoside, which also exhibits a complex absorbance spectrum with peaks at 287, 298 and 307 nm, in addition to the main bands given in Table 3.1. These extra transitions may be due to vibrational structure which is not resolved in the other compounds.

7-Phenyl-7-deazaadenosine has a similar λ_{max} to an equivalent 8-modified nucleoside, 8-phenyl-2'-deoxyadenosine, which absorbs at 289 nm in DMSO.[91] Their molar absorption coefficients are also similar.   However, the differences between 7-phenylethynyl-7-deazaadenosine and 8-phenylethynyladenosine are more pronounced. In this case, the 7-modified compound has a much shorter absorbance wavelength than the 8-modified compound (302 nm and 330 nm respectively).[51]   Investigations into the effect of substitution patterns in phenylacetylene dendrimers[147] and bis(phenylacetylene)benzene systems[148] have shown that *meta* substitution of the phenyl ring leads to interruption of the conjugation and isolation of one phenylacetylene unit from the other. It has been observed that 1,3-di(phenylethynyl)benzene has a UV spectrum similar to phenylacetyene, however 1,2- and 1,4-di(phenylethynyl)benzene have UV spectra consistent with an extended π-system. Moving the phenylacetylene substituents from the 8-position in 8-(phenylethynyl)adensoine, to the 7-position in 7-(phenylethynyl)-7-deazaadenosine has a

similar effect, disrupting conjugation through the purine ring. The 8-modified nucleoside therefore has a larger conjugated system because it includes the pyrimidine ring (Figure 3.4).

Table 3.1 UV-Vis spectra of modified nucleosides in DMSO.
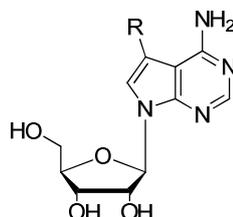
	R		λ_{max} /nm	$\lambda_{shoulder}$ /nm	ε /mol^{-1}dm^{3}cm^{-1}
1		**16**	302	322	18300
2		**17**	337	353	12900
3		**18**	286	-	12400
4		**20a**	321	297	21500
5		**20b**	321	306	22600
6		**20c**	331	309	25400
7		**20d**	313	315	21400
8		**20e**	328	321	14900

The absorbance spectrum of diphenylacetylene in DMSO has not been reported, however its spectra in a variety of other solvents (such as acetonitrile,[149] dichloromethane,[150] cyclohexane and ethanol [151]) are very different from those observed for the diphenylacetylene-functionalised nucleosides. In ethanol, diphenylacetylene has peaks at 298, 289, 280, 273 and 265 nm.[151] The majority of these peaks are due to vibrational transitions, and all the peaks are significantly blue-shifted with respect to the absorbances observed for the diphenylacetylene-decorated nucleosides. This suggests that rather than

the diphenylacetylene moiety acting as an extrinsic fluorescent label, the nucleobase is extending the conjugation and is part of the chromophore. This is further evidence that, as shown in Figure 3.4, the nucleobase is part of the chromophore.
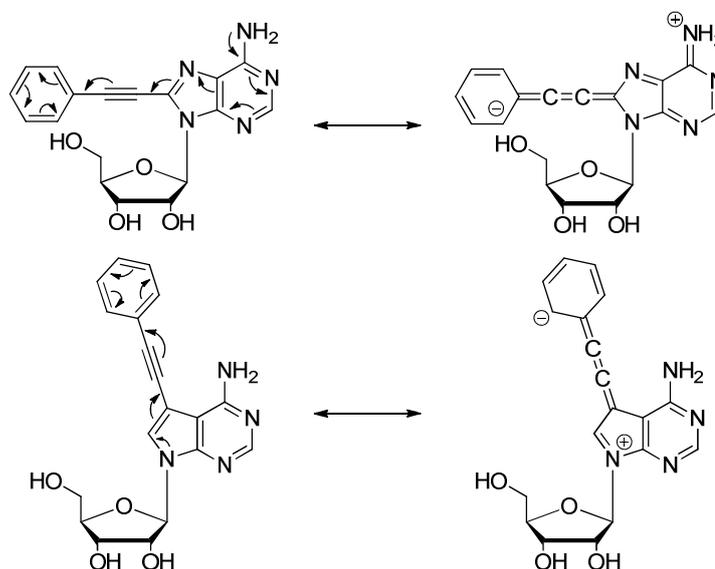
Figure 3.4 Differences in conjugation between 8-(phenylethynyl)adenosine and 7-(phenylethynyl)-7-deazaadenosine.

In general, changing the substituent on the terminal position of the diphenylacetylene moiety appears to have a relatively small effect on the UV spectrum. Examples with electron-donating (MeO- and MeS-, Table 3.1, entries 5 and 6) and electron-withdrawing (F_3C-, entry 8) substituents were synthesised; however there does not appear to be any pattern to the UV absorbance maxima. This may be because the heterocycle can act as an electron-donating or electron-withdrawing group, depending on the substituents on the chromophore. This pattern was previously observed in both the 8-phenylethynyl- and 8-phenyladenosine series.[152] A similar effect was also reported for 2,5-bis(phenylethynyl)thiophenes.[153] However, in this case, there are not enough examples to definitively know if this pattern is being followed.

The 3-thienylethynylphenyl-substituted nucleoside was investigated because 3-thienylethynylphenyl-substituted heterocycles have shown bathochromically shifted UV and emission spectra, compared with phenylethynylphenyl-substituted analogues.[154] The

thienylethynylphenyl-substituted nucleoside, however, did not show red-shifted UV absorbance.

Overall, the modified nucleosides showed promising UV spectral properties. The 4-(phenylethynyl)phenyl-subsituted 7-deazaadenosines have absorbance wavelengths above 320 nm, suitably shifted away from the natural nucleic acid absorbances (260 nm) and tryptophan residue absorbances (280 nm). Their molar absorption coefficients are also high (*ca.* $20 \times 10^3 \text{ mol}^{-1}\text{dm}^3\text{cm}^{-1}$) suggesting they are effective chromophores.

3.3   Fluorescence spectroscopy

The fluorescence emission spectra of the modified nucleosides were also recorded in DMSO. The spectra were recorded at concentrations which give a UV absorbance of *ca.* 0.1 A.U., in order to minimise inner filter effects. Quantum yields were measured using an integrating sphere. An integrating sphere reflects >99.9% of the photons emitted and allows a steady-state fluorimeter to be used for a direct (primary) measurement of quantum yield. It is also much faster and more efficient than the traditional ratiometric (secondary) method.[155] The integrating sphere and method used is similar to that described by Beeby and co-workers.[156] Measurements are made to determine the number of photons absorbed by the sample (after subtraction of the background solvent) by measuring the emission over a range at the excitation wavelength ±10 nm. The number of photons emitted is then measured by subtracting the emission spectrum of the background solvent from the emission spectrum of the sample. These numbers can be used to give a ratio of photons absorbed to photons emitted, provided the spectrometer response is accounted for.

Unlike the analogous 8-modified adenosines, the 7-modified-7-deazaadenosines with small substitutions (Table 3.2, entries 1-3) show only weak fluorescence. As described in Section 3.3, this is probably due to the reduced conjugation when the substituent is on the 7-position. These nucleosides had low quantum yields (less than 2%), compared with the high quantum yields seen for 8-phenyl- and 8-phenylethynyladenosines (0.81 and 0.55 respectively) in DMSO. The arylethynylphenyl-labelled nucleosides, however, show more promising fluorescent properties.

Table 3.2 Fluorescence spectroscopy of the modified nucleosides in DMSO.
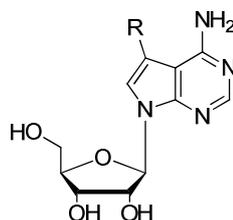
	R		λ_{em} /nm	Stokes' shift /cm^{-1}	Φ	τ /ns
1		**16**	weak	-	<0.01	4.1 (87%), 0.5 (13%)
2		**17**	432	6438	0.02	1.5 (60%), 0.5 (40%)
3		**18**	363	7663	<0.01	n.d.[†]
4		**20a**	428	7788	0.74	2.2 (92%), 0.5 (8%)
5		**20b**	405	6461	0.78	1.6 (90%), 0.5 (10%)
6		**20c**	431	7009	0.76	2.0 (92%), 0.5 (8%)
7		**20d**	404	7196	0.32	1.5 (88%), 0.5 (12%)
8		**20e**	491	10121	0.41	2.1 (87%), 0.8 (12%)

All the arylethynylphenyl-substituted analogues have high quantum yields, in particular compounds **20a**, **20b** and **20c** (Table 3.2, entries 4-6) which all have quantum yields over 70%. The thiophene-containing analogue **20d** (entry 7) had a lower quantum yield (32%), as did the CF$_3$-compound **20e** (entry 8, 41%). The variation in the wavelength of the emission maxima for these compounds was much larger than the variation in their absorbance maxima. This gives rise to a large variation in Stokes' shift. In particular, the CF$_3$-compound has a large Stokes' shift of 10x10^3 cm^{-1}. Although fluorescence emission spectra generally mirror the absorbance spectra of the fluorophore, in this case the emission spectra have only one peak, and do not exhibit the shoulder peaks seen in the

---

[†] A good fit could not be obtained for the data collected (see Appendix I).

absorbance spectra. This means that although there are two electronic absorbance transitions, there is only one fluorescence emission transition. There are two possible explanations for this: either only one absorbance is emissive, or one of the excited states converts into the other, before emission from this excited state.

Fluorescence lifetimes were measured in collaboration with Dr. A Beeby (Durham) using Time-Correlated Single Photon Counting (TCSPC). All the 7-modified-7-deazaadenosines were found to have bi-exponential decays, with a major component which varied according to the substituent at the 7-position, and a minor component (ca. 10%) which was generally around 0.5 ns. A representative decay spectrum is shown in Figure 3.5. The fluorescence lifetimes of the deoxyribo-analogue **35** (2.3 ns (88%), 0.6 ns (12%)) were very similar to those measured for **20a**. Changing the solvent from DMSO to water gave quite different lifetime measurements for **35**. Like in DMSO, a double exponential decay gave the best fit, but in this case the shorter lifetime component is the dominant pathway (0.4 ns (62%), 0.9 ns (38%)), and both lifetimes are significantly shorter than the 2.3 ns measured in DMSO. The presence of two lifetimes could be due to the fluorophore being in two different environments – this could be indicative of aggregation, although no other evidence for this was observed. Alternatively, it could be due to the presence of a second fluorophore, as simple substitutions at C7 (such as iodide) give rise to weak fluorescence emission from the heterocycle. In this case, the disconnection between the aromatic systems gives rise to two separate fluorescent emissions (Figure 3.4).
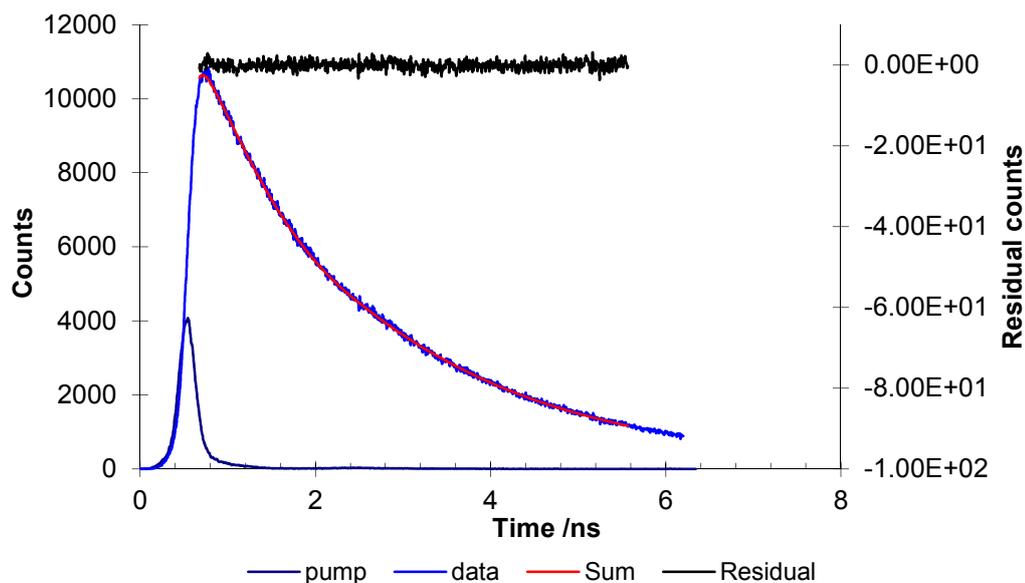
Figure 3.5 Fluorescence decay spectrum of **20a** in DMSO.

3.4 Solvatochromism

Solvatochromism of fluorophores can be informative about the nature of the excited state and the electronic transitions giving rise in the absorbance and emission spectra. The effects of solvent on the absorbance and emission spectra of 7-[4-(phenylethynyl)-phenyl]-7-deazaadenosine (**20a**) were therefore investigated (Table 3.3). Solvatochromism also has implications for biological applications – solvatochromic fluorescent probes can be used to gain insight into the polarity of the local environment. As water is the "biological solvent", the way the fluorophore behaves in aqueous conditions is very important. These results show that there is a significant drop-off in quantum yield in water, as compared to DMSO. Although this is not ideal, the quantum yield of 15% is still useful,[157] and is an order of magnitude better than the analogues with simple phenyl- or phenylethynyl- substitutions, as well as other reported compounds of this type.[27]

Table 3.3 The effect of solvent on the spectroscopic properties of **20a**.

Solvent	λ_{max} /nm	ε /mol^{-1}dm^{3}cm^{-1}	λ_{em} /nm	Stokes' shift /cm^{-1}	Quantum Yield
DMSO	321	21531	428	7788	0.74
H$_2$O[a]	302	19764	429	9802	0.15
MeOH	306	24875	407	8110	0.46
CH$_2$Cl$_2$[a]	312	25286	394	6670	0.61
MeCN	308	24313	409	8017	0.62
EtOH	304	26312	393	7449	0.31
^{i}PrOH	304	31101	396	7642	0.29
EtOAc[a]	311	23944	389	6447	0.33
DMF	318	26630	420	7637	0.78

[a] Solutions made from stocks of **20a** in DMSO.

There are relatively small changes in the absorbance and emission spectra for the modified nucleoside as the solvent polarity changes. The absorbance maximum varies between 302 nm (water) and 321 nm (DMSO), while the emission maximum varies between 393 nm (ethanol) and 428 nm (DMSO). When these changes in absorbance and emission maxima are plotted against solvent dielectric, there appears to be only a weak correlation (Figure 3.6). Other measures of solvent polarity, such as E_T(30),[158] also give very weak correlations (Figure 3.7). If the transition seen is $\pi \rightarrow \pi^*$, very little change in the λ_{max} should be observed on changing the solvent, however if the UV band is due to intramolecular charge transfer (ICT), a significant bathochromic shift would be expected.[145] In this case, there is no clear correlation between the absorbance wavelength and the solvent polarity, suggesting that the observed UV transitions are not due to ICT. There may be other competing factors, such as preferential solvation of the ground or excited electronic states.

Although there are only poor correlations between the absorbance and emission wavelengths, the results do show a strong correlation between the Stokes' shift and the solvent dielectric (Figure 3.8). This shows that the fluorophore is sensitive to its environment, and can detect changes in its local polarity.
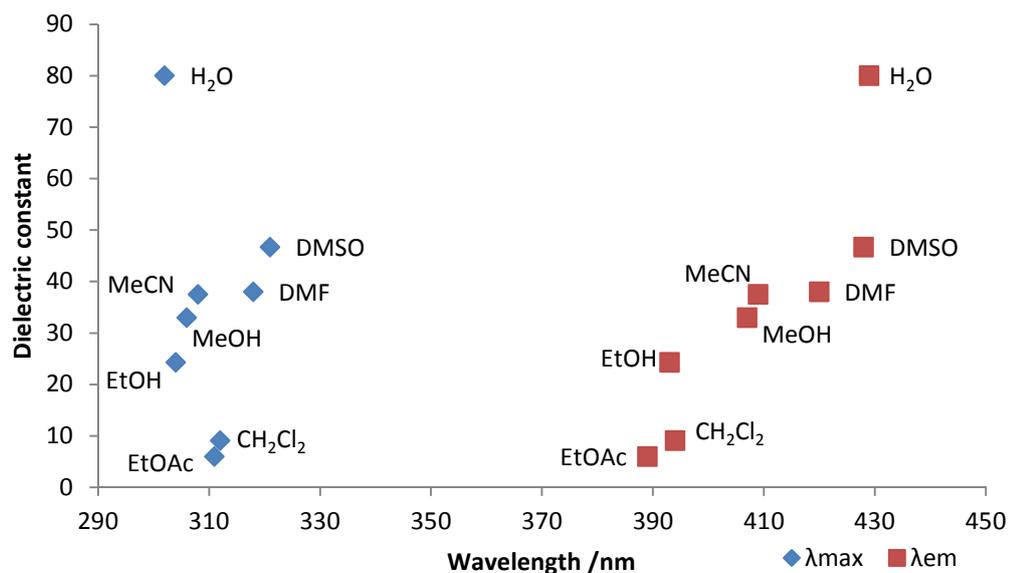
Figure 3.6. Changes in λ_{max} and λ_{em} against solvent dielectric for **20a**.
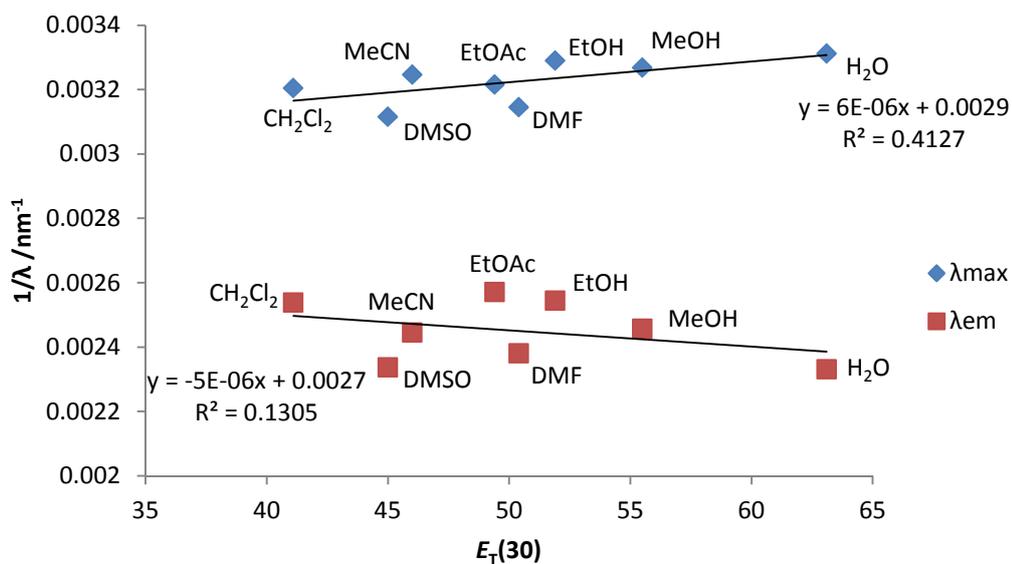
Figure 3.7 Correlation between $E_T(30)$ and wavenumber for **20a**.
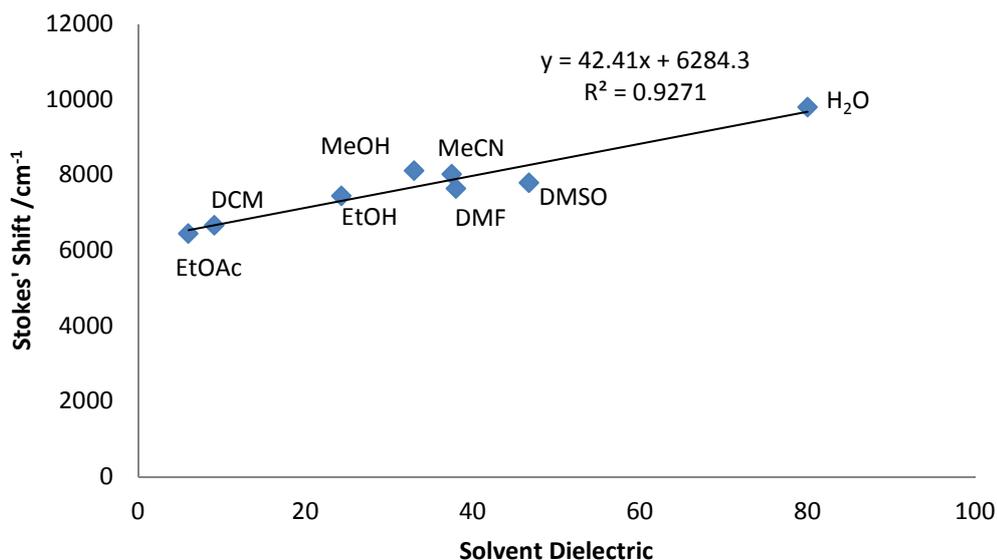
Figure 3.8 Correlation between solvent dielectric and Stokes shift for **20a**.

3.5 Effect of pH

The effect of changing pH on the UV-Vis and fluorescence spectra can also be informative about the nature of the electronic transitions observed.[159] To probe the observed spectra further, the effect of pH on the UV-Vis spectra has been investigated for **20a** in DMSO by addition of aqueous 1 M HCl (Figure 3.9). It has been previously shown that addition of acid to a compound exhibiting ICT bands has a dramatic effect on the UV-Vis spectrum.[160] This is generally due to protonation of a lone pair which is then no longer able to donate into the ICT system. In this case, it does appear that a similar, if less pronounced, effect is seen. The main band and the shoulder peak appear to give an "averaged" peak upon addition of acid. Although this could point to the presence of an ICT band it is not conclusive. Rosemeyer and Seela have previously shown by ^{15}N NMR spectroscopy that protonation of 7-deazaadenosine nucleosides occurs at the N-1 position.[161] They also noted that the 7-deaza-nucleoside was more basic than adenosine.
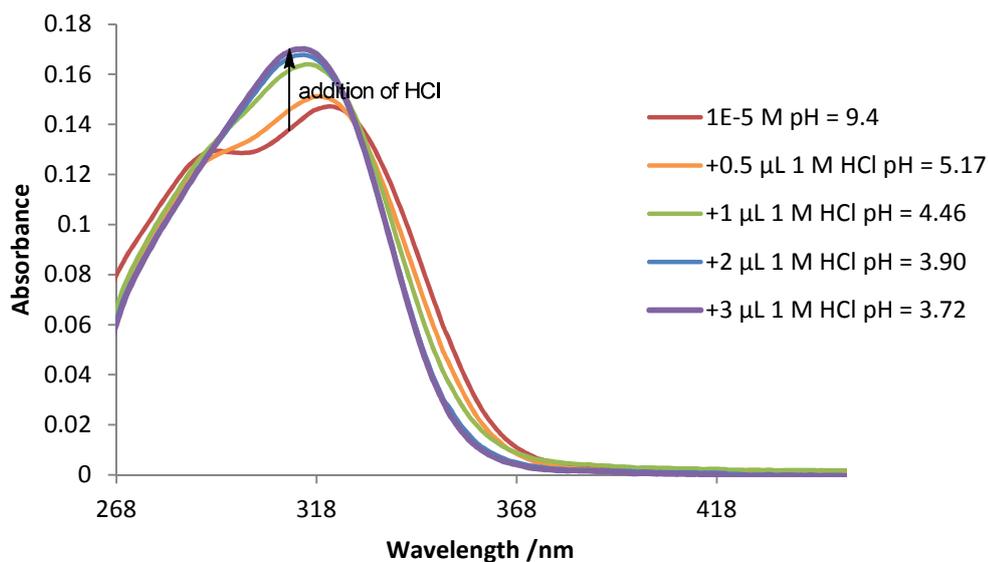
Figure 3.9 Effects of decreasing pH on **20a** in DMSO.

This initial result shows a clear change in the UV spectrum upon protonation. To investigate this further, buffered solutions were used to probe the effect of pH on the fluorescence emission spectra of **20a** and **35**. The fluorescence emission spectrum showed pronounced quenching on protonation of the nucleoside. As protonation is known to occur on the N-1 position, this suggests that the pyrimidine heterocycle is heavily involved in the fluorescence emission. The dependence of the fluorescence spectra of **35** on pH is shown in Figure 3.10. The intensity at the emission maximum was measured as a function of pH and displayed a sigmoidal transition indicative of a single ionisable group being titrated. This pH titration curve was fitted with a sigmoidal function (using OriginLab) to determine the pK_a. The pK_a was found to be 4.35, which is 0.95 pH units less than 2'-deoxytubercidin, and 0.85 pH units more than 2'-deoxyadenosine. The difference in pK_a between the substituted and unsubstituted 7-deazaadenosines has been previously observed for other 7-modified 7-deazaadenosines. 7-halo-7-deazadenosines have pK_a values between 4.24 (7-bromo) and 4.42 (7-fluoro), measured by UV-Vis spectroscopy.[162] This is thought to be due to the electron-withdrawing effect of the halogens, however there does not appear to be a direct correlation between the electronegativity of the substituent and the pK_a of the heterocycle. 7-(1-Propynyl)-7-deazaadenosine also has a lower pK_a than the unsubstituted heterocycle (pK_a = 4.50), which is quite similar to the pK_a measured for **35**.
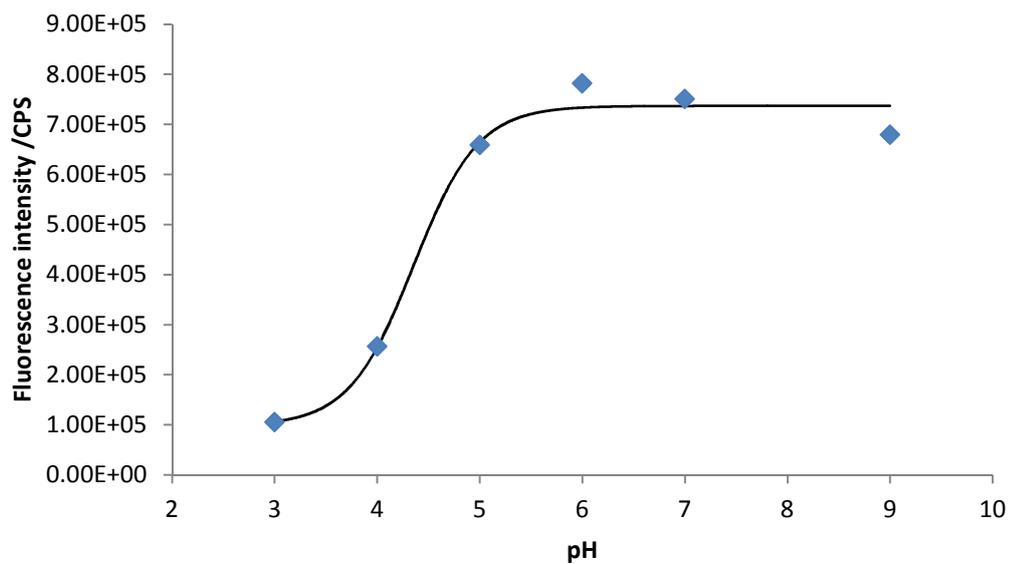
Figure 3.10 Changes in fluorescence intensity of compound **35** at λ_{429} in buffered solution with varying pH at 2.5 µM.

3.6 *Ab-initio* calculations

In an attempt to better understand the differences in the photophysical properties of the first- and second-generation 7-modified-7-deazaadenosines, some calculations on model systems (*N*-methyl-7-modified-7-deazaadenines) were carried out by Dr J. Slattery (Department of Chemistry, University of York).
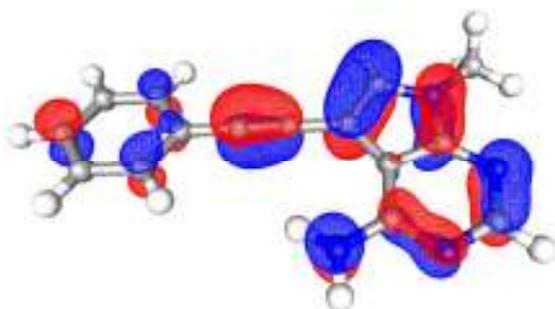
HOMO: -4.98 eV   HOMO: -5.02 eV
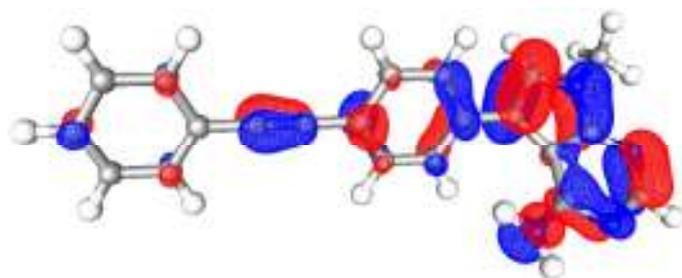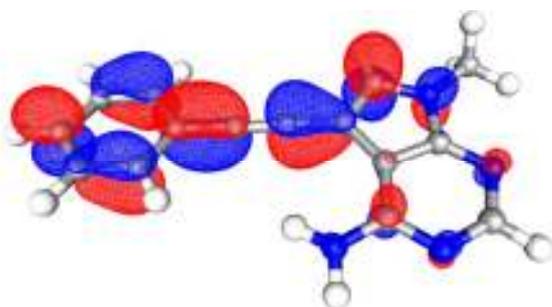
LUMO: -2.25 eV   LUMO: -2.53 eV
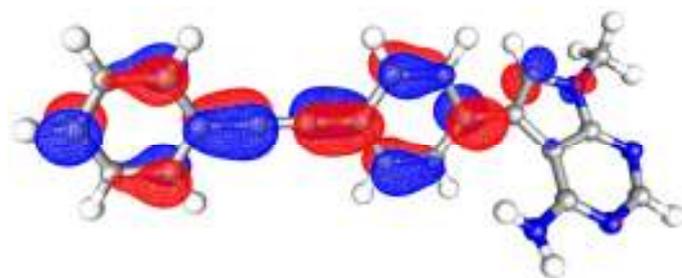
Figure 3.11 BP86/def2-TZVPP optimised structures for N-methyl-7-phenylethynyl-7-deazaadenine (left) and N-methyl-7-[4-(phenylethynyl)-phenyl]-7-deazaadenine (right).

The calculations suggest that the HOMO and LUMO of *N*-methyl-7-phenylethynyl-7-deazaadenine are distributed quite evenly across the compound, with the HOMO slightly more on the nucleobase, and the LUMO weighted towards the phenylethynyl substituent (Figure 3.11). Conversely, the HOMO and LUMO of *N*-methyl-7-[4-(phenylethynyl)-phenyl]-7-deazaadenine show a greater separation, with the HOMO on the nucleobase, and the LUMO predominantly on the diphenylacetylene subsituent. This suggests that the modified nucleoside could act as a donor-acceptor system, indicating ICT is theoretically possible.

Although the wavelengths of the transitions determined from the *ab initio* calculations do not match the measured UV spectra, the calculated values are qualitatively similar. The phenylethynyl- system was found to have the higher energy transition, as observed experimentally.

3.7   Conclusions

The photophysical properties of the 7-modified-7-deazaadenosines with simple phenyl- or arylethynyl- groups were found to be insufficient for applications in biological systems, mainly due to their low quantum yields. However, the novel nucleosides with arylethynylaryl-moieties exhibited promising UV and fluorescence properties, with the UV absorbance maxima shifted away from the regions associated with the aromatic amino acid residues. The quantum yields were high in DMSO, and although the quantum yield of the parent compound decreased in water (to 15%), this is still an order of magnitude more than simple derivatives, such as the phenyl- and phenylethynyl-substituted compounds. The changes in the absorbance and emission wavelengths on changing the solvent were not consistent with a classical ICT push-pull system, however the Stokes' shift was found to be sensitive to changes in polarity. Although there is not a clear correlation between local polarity and the absorbance and emission maxima, the Stokes' shift correlations may be useful as a probe of local environment in biological systems.

4   Oligonucleotide properties

4.1   Introduction

Incorporation of modified nucleotides into nucleic acids can result in a variety of changes, to both the properties of the nucleic acid, and the properties of the modified nucleotide. For fluorescent nucleotides, this can mean that the fluorescent modification can impact on the structure of the nucleic acid. Many fluorescent base-modifications can affect Watson-Crick base-pairing, for adenine nucleotides this generally means modifications to the exocyclic amine or the C2 position. Disruption to Watson-Crick base-pairing will decrease duplex stability, particularly for shorter oligonucleotides. Other base-modifications, such as 8-modified purines, have been shown to interfere with duplex stability by inducing changes in sugar conformation.[163] It is equally as important to know what impact the environment of the nucleic acid has on the modification which has been introduced. This is particularly important for fluorescent nucleotides, which often show significant changes upon incorporation. For example, 2-aminopurine, a commonly used fluorescent nucleobase, shows significant quenching (about 2 orders of magnitude) upon incorporation into nucleic acids.[164]

A variety of techniques can be used to determine the impact of the modified nucleotide on the oligonucleotide structure in both the single strands and the duplexes. Changes in the UV/Vis and fluorescence spectra can be informative, particularly for fluorophores sensitive to microenvironment.

4.1.1   Melting temperatures

Duplex destabilisation is generally measured by comparing the melting temperature of the duplex containing the modification with the unmodified duplex. This gives an indication of whether the modified nucleotide is forming a stable base-pair. DNA shows a hyperchromic change in absorbance upon melting (denaturing). When the DNA duplex is heated slowly, an increase in UV absorbance is observed (usually measured at 260 nm). The transition observed is cooperative; as the duplex begins to melt, the melting becomes more facile. This is shown by the characteristic sigmoidal melting curve (Figure 4.1). The melting

temperature (T_m) is defined as the temperature at which half the duplex has melted. It is measured as the point half way between the minimum and maximum UV absorbances. Alternatively, the melting curve can be differentiated to give a characteristic first-derivative curve, where the melting temperature is the temperature at the maximum of the curve (the maximum gradient of the melting curve and mid-point of the melting transition). Melting temperatures can also be used to assess base-discrimination with mismatched base pairs, by examining the change in melting temperature when a base mismatch is present. Usually, a significant decrease in melting temperature is observed when a base-pair mismatch is present as compared to the fully complementary sequence. Many frequently used methods for identification of SNPs rely on changes in melting temperatures, but using techniques which can differentiate between double-stranded and single-stranded DNA.
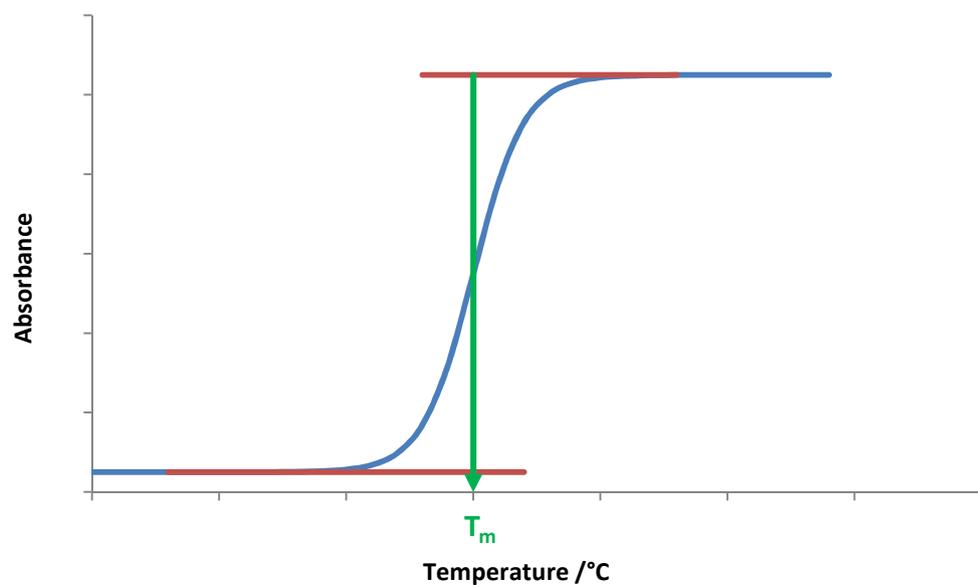
Figure 4.1 Representative sigmoidal melting curve.

4.1.2 Circular dichroism spectroscopy

Circular dichroism (CD) is a useful technique for discerning information about the secondary structure of chiral biomolecules such as proteins and nucleic acids. This technique measures the difference between a sample's absorbance of left and right circularly polarised light, which is dependent on the secondary structure of the nucleic acid.[165] Circular dichroism absorbances are generally converted to extinction coefficients,

to give a concentration-independent value. For biopolymers, the value is generally calculated as a mean residue concentration. Another commonly used unit for circular dichroism is molar ellipicity ($[\theta]$), which is directly related to the molar circular dichroism ($\Delta\varepsilon$) (Equation 4.1).

$$\Delta\varepsilon = \varepsilon_{LCP} - \varepsilon_{RCP} = \frac{[\theta]}{3298.2}$$

Equation 4.1 Relationship between molar circular dichroism and molar ellipticity.

It is important to determine the effect of introducing modified nucleotides has on the secondary structure, and circular dichroism is an important tool for this. The different nucleic acid secondary structures, A-form, B-form and Z-form, each have characteristic circular dichroism spectra. CD spectra of oligonucleotides have contributions from two effects: the base-sugar interaction (dependent on the sugar conformation), and base-base interactions (dependent on stacking effects).

The key circular dichroism features of B-form DNA can be seen in a representative spectrum of calf thymus DNA (Figure 4.2). Characteristically, B-form DNA exhibits a positive peak at 270-280 nm and a negative peak at 245 nm which can be clearly seen in this typical spectrum. The positive peak at 220 nm is also characteristic, but closer to 200 nm the voltage of the detector begins to increase, making this region less accurate.
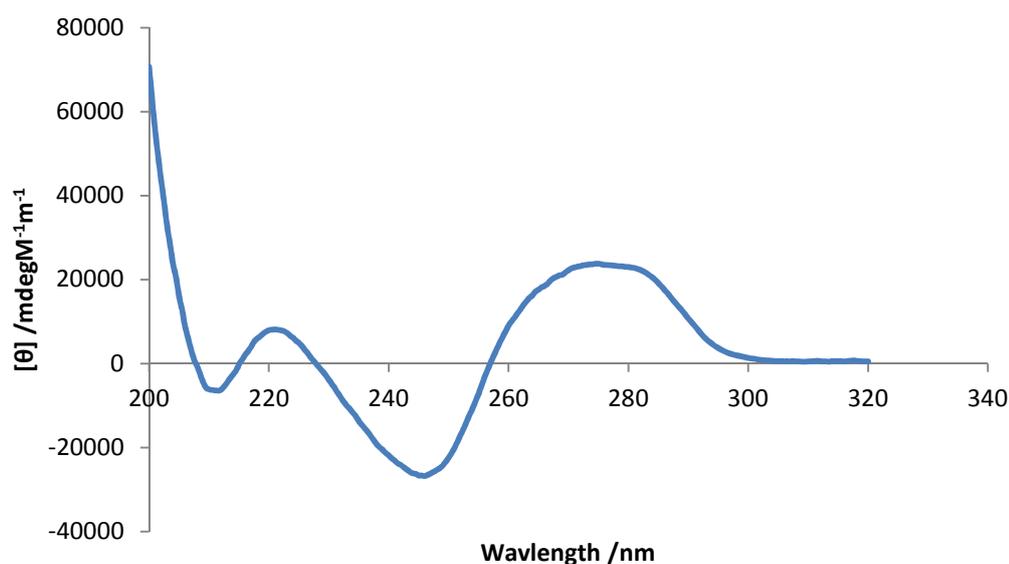
Figure 4.2 CD spectrum of calf thymus DNA.

4.1.3 Fluorescence quenching experiments

Fluorescence quenching experiments can be used to investigate the local environment of a fluorophore, specifically to provide information about solvent accessibility or local charge. For example, quenching experiments are frequently used to determine the environment of tryptophan residues within a protein.[166] There are a variety of mechanisms by which a quencher molecule can interact with fluorophores. The mechanisms can be divided into two main categories: static quenching and dynamic quenching. Static quenching involves formation of a non-fluorescent complex between the fluorophore and a quencher, whereas dynamic quenching requires a collision between the quencher and the fluorophore which quenches the excited state of the fluorophore in a non-radiative process. The kinetics of quenching reactions can therefore help to distinguish the two types of quenching. Dynamic (collisional) quenching is described by the Stern-Volmer equation (Equation 4.2).[167]

$$\frac{F_0}{F} = \frac{\tau_0}{\tau} = 1 + k_q\tau_0[Q] = 1 + K_{SV}[Q]$$

Equation 4.2 Stern Volmer equation for collisional quenching.

F_0 = Fluorescence intensity in the absence of a quencher; F = Fluorescence intensity in the presence of a quencher; k_q = bimolecular quenching constant; τ_0 = lifetime in the absence of quencher; τ = lifetime in the presence of quencher; [Q] = concentration of quencher; K_{SV} = Stern-Volmer quenching constant.

Therefore, for dynamic quenching, plots of F_0/F against [Q] should give linear plots; however linear plots are also observed for static quenching. The different mechanisms can be identified by looking at the effect of temperature on fluorescence quenching in the experimental system. Generally, increasing the temperature will cause increased collisional (dynamic) quenching. Static quenching is reduced at higher temperatures due to dissociation of the quencher-fluorophore complex. Systems in which both static and dynamic quenching occur often display upward curving plots of F_0/F against [Q]. This data can be described by the exponential function given in Equation 4.3.[168]

$$\frac{F_0}{F} = (1 + K_{SV}[Q])e^{V[Q]}$$

Equation 4.3 Stern-Volmer equation describing both static and dynamic quenching.

V = static quenching constant.

A variety of quenchers can be used, depending on the application. Compounds containing heavy atoms can quench fluorescence through spin-orbit coupling. For these applications, caesium or iodine are often used. Caesium is usually added as caesium chloride, which has been shown to be only a moderately efficient quencher of tryptophan (Table 4.1). It is also very large, limiting its access to small solvent cavities. As it is positively charged, it is not likely to quench fluorophores in positively charged or hydrophobic environments. Iodide, usually as potassium iodide, has many of the same disadvantages as caesium when used as a quencher, although it has been shown to be more efficient for quenching tryptophan residues. Potassium iodide is also light-sensitive, so is usually stabilised with 2-mercaptoethanol or sodium thiosulfate. In systems compatible with organic solvents, organoiodides such as iodobenzene can be used as neutral iodide quenchers.

A highly efficient (and water-soluble) neutral fluorescence quencher for tryptophan is acrylamide (prop-2-enamide). Acrylamide has very high permeability of hydrophobic pockets in proteins so can quench tryptophan residues not normally accessible to caesium chloride or potassium iodide.[169]

Table 4.1 Efficiency of indole fluorescence quenching at pH 7.0 for several quenchers.[166]

Quencher	k_q /10^{-9}M^{-1}s^{-1}	V /M^{-1}
Cs$^+$	1.1	-
I$^-$	6.4	2.0
acrylamide	7.1	2.5

4.1.4 Application to detection of single-nucleotide polymorphisms

The modified nucleotide contains a fluorophore which is in conjugation with the nucleobase. This means it should be directly sensitive to formation of a Watson-Crick base-pair (Figure 4.3). Photophysical investigations with the nucleoside have shown that the fluorophore shows some environmental sensitivity, so it may be able to detect changes in local polarity and solvent accessibility. Therefore, the nucleotide should be an ideal probe of base-pair mismatches. As described in Chapter 1, base-pair mismatches arising from single nucleotide polymorphisms are associated with susceptibility to many diseases.

Therefore, simple fluorescence detection of the presence or absence of a SNP is a valuable tool for identifying genetic variations. The effect of base-pair mismatches on the fluorescence properties of the oligonucleotides, as well as any structural or physical changes, are key to understanding base-pair discrimination. This will dictate the usefulness of the modified nucleotide for assaying SNPs.
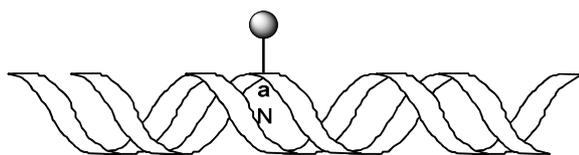
Figure 4.3 Investigation of base-pair mismatches using the modified oligonucleotides.

4.1.5   Oligonucleotide sequences

The sequences of the modified oligonucleotides and their complementary strands used in this section are given below. The sequences were chosen based on those used by Saito and co-workers for investigating base-pair mismatch discrimination. The range of sequences chosen reflects a wide range of environments for the modified nucleotide. The bases flanking the modified nucleotide are varied between C, T and G. Adenine nucleotides are not used as flanking bases to avoid structural distortions which can be caused by runs of A.[170] The bases in the rest of the sequence are also varied to give a range of sequence types, including G·C rich and both purine-rich and pyrimidine-rich sequences. Many base-discriminating nucleotides are only tested in one or two different sequences, but examination of a variety of sequences is vital to determine any sequence dependency or specific quenching effects.

For oligonucleotides **C1**, **C2**, **C3**, **C4** and **C5**, sequences are as **ODN1**, **ODN2**, **ODN3**, **ODN4** and **ODN5**, with **a** = A.

ODN1	$^{5'}$d(CGCTTG**a**GTTCGC)$^{3'}$
1T	$^{3'}$d(GCGAACTCAAGCG)$^{5'}$
1C	$^{3'}$d(GCGAACCCAAGCG)$^{5'}$
1A	$^{3'}$d(GCGAACACAAGCG)$^{5'}$
1G	$^{3'}$d(GCGAACGCAAGCG)$^{5'}$
ODN2	$^{5'}$d(CGCTTC**a**CTTCGC)$^{3'}$
2T	$^{3'}$d(GCGAAGTGAAGCG)$^{5'}$
2C	$^{3'}$d(GCGAAGCGAAGCG)$^{5'}$
2A	$^{3'}$d(GCGAAGAGAAGCG)$^{5'}$
2G	$^{3'}$d(GCGAAGGGAAGCG)$^{5'}$
ODN3	$^{5'}$d(CGCAAG**a**GAACGC)$^{3'}$
3T	$^{3'}$d(GCGTTCTCTTGCG)$^{5'}$
3C	$^{3'}$d(GCGTTCCCTTGCG)$^{5'}$
3A	$^{3'}$d(GCGTTCACTTGCG)$^{5'}$
3G	$^{3'}$d(GCGTTCGCTTGCG)$^{5'}$
ODN4	$^{5'}$d(CGCAAC**a**CAACGC)$^{3'}$
4T	$^{3'}$d(GCGTTGTGTTGCG)$^{5'}$
4C	$^{3'}$d(GCGTTGCGTTGCG)$^{5'}$
4A	$^{3'}$d(GCGTTGAGTTGCG)$^{5'}$
4G	$^{3'}$d(GCGTTGGGTTGCG)$^{5'}$
ODN5	$^{5'}$d(CGCAAT**a**TAACGC)$^{3'}$
5T	$^{3'}$d(GCGTTATATTGCG)$^{5'}$
5C	$^{3'}$d(GCGTTACATTGCG)$^{5'}$
5A	$^{3'}$d(GCGTTAAATTGCG)$^{5'}$
5G	$^{3'}$d(GCGTTAGATTGCG)$^{5'}$

4.2   Spectroscopic properties of the modified oligonucleotides

The UV-Vis and fluorescence properties of the single-stranded oligonucleotides were found to be significantly different from the nucleoside monomer. The absorbance spectra for the oligonucleotides show a peak at 322 nm, a 20 nm bathochromic shift compared to the nucleoside in water. The change in absorbance wavelength means that the absorbance spectrum is no longer obscured by the intrinsic absorbances of nucleic acids. This change was observed for all the oligonucleotides synthesised (Figure 4.4). The fluorescence emission maxima were also shifted, from 428 nm to *ca.* 400 nm, with only small variation between sequences (Table 4.2, Figure 4.5). The molar absorbance coefficients were larger

than for the nucleoside, and the Stokes' shifts were smaller, again with only small variations between sequences. The quantum yields, however, varied from 0.024 (**ODN2**) to 0.237 (**ODN3**), depending on the sequence. The origins of this variation are not clear, as comparing **ODN1** and **ODN3**, which both have the modified adenosine adjacent to guanine, differ in quantum yield by an order of magnitude (0.025 and 0.237, respectively). **ODN2** and **ODN4**, which have the modified base adjacent to cytosine, also differ significantly (2.4% and 8.6% respectively).
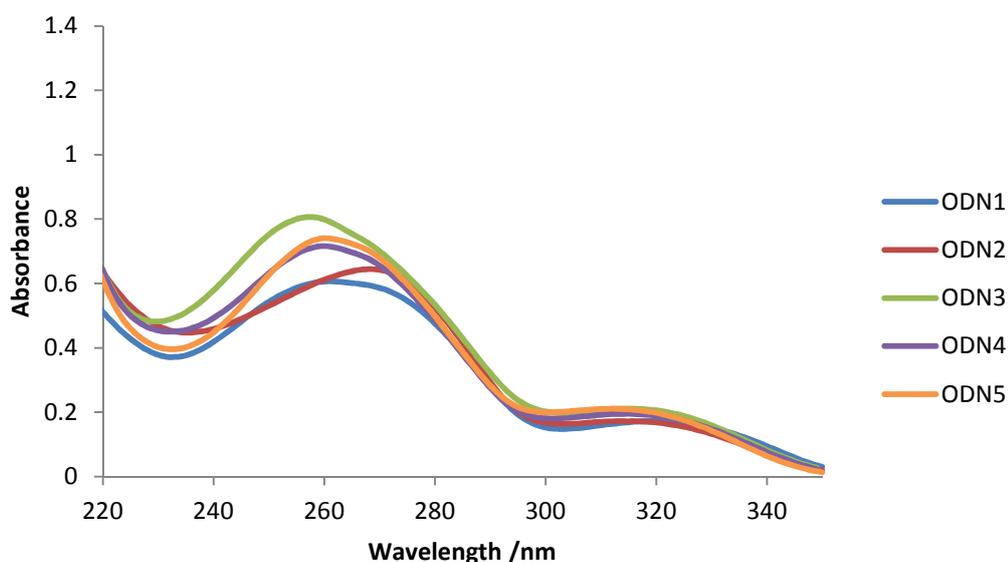
Figure 4.4 Absorbance spectra of modified oligonucleotides (5 µM).

The fluorescence spectra change significantly upon formation of a duplex (Table 4.3). In most cases, the emission maximum wavelength increases slightly (6-12 nm), however for **ODN5**, the emission wavelength decreases by 5 nm. This will be discussed further in Section 4.4.1. **ODN1** and **ODN2** show increases in quantum yield upon formation of the duplex, but **ODN3**, **ODN4** and **ODN5** show decreases in quantum yield compared to the single-strand oligonucleotides. In the duplex, the quantum yields appear to have some dependency on the sequence. **ODN1** and **ODN3** have the highest quantum yields (11.0% and 15.4 % respectively), and both have the modified nucleotide located between guanines. The other oligonucleotides have significantly lower quantum yields, and all have the modified nucleotide adjacent to pyrimidines.

Table 4.2 UV and fluorescence properties of the single-strand oligonucleotides containing the modified base.

	Sequence	$\lambda_{em}{}^{a}$ /nm	Stokes' shift /cm^{-1}	Φ
ODN1	$^{5'}$d(CGCTTG**a**GTTCGC)$^{3'}$	400	6056	0.025
ODN2	$^{5'}$d(CGCTTC**a**CTTCGC)$^{3'}$	400	6056	0.024
ODN3	$^{5'}$d(CGCAAG**a**GAACGC)$^{3'}$	405	6365	0.237
ODN4	$^{5'}$d(CGCAAC**a**CAACGC)$^{3'}$	404	6303	0.086
ODN5	$^{5'}$d(CGCAAT**a**TAACGC)$^{3'}$	405	6365	0.061

^{a} λ_{ex} = 322 nm.
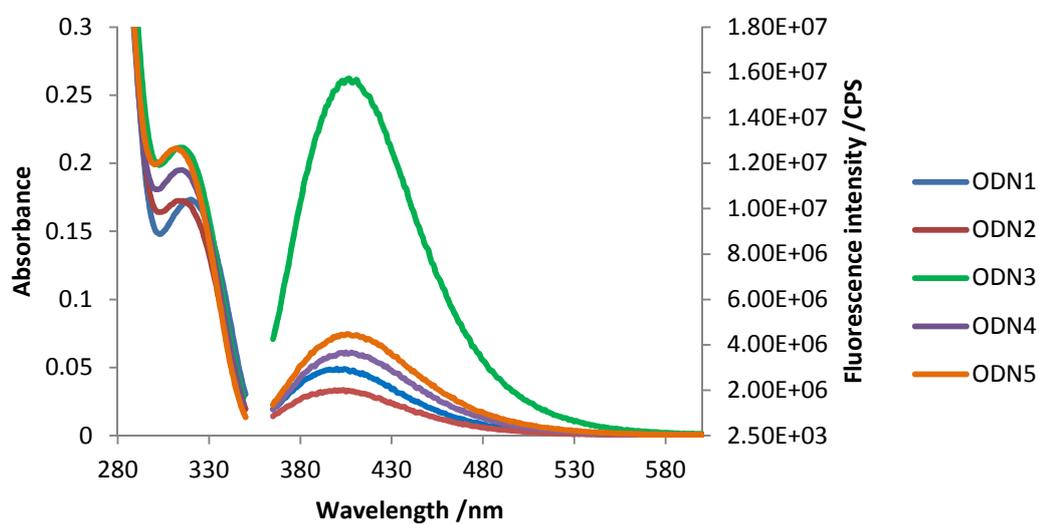
Figure 4.5 Absorbance and emission spectra of modified oligonucleotides.

Table 4.3 Comparison of emission wavelength maxima and quantum yields between single-strand oligonucleotides and their duplexes.

	Sequence	λ_{em} dsDNA[a] /nm	$\Delta\lambda_{em}$[b] /nm	Φ dsDNA	$\Delta\Phi$[c]
ODN1	$^{5'}$d(CGCTTG**a**GTTCGC)$^{3'}$	412	+12	0.110	+0.085
ODN2	$^{5'}$d(CGCTTC**a**CTTCGC)$^{3'}$	409	+9	0.035	+0.011
ODN3	$^{5'}$d(CGCAAG**a**GAACGC)$^{3'}$	412	+7	0.154	-0.083
ODN4	$^{5'}$d(CGCAAC**a**CAACGC)$^{3'}$	410	+6	0.054	-0.034
ODN5	$^{5'}$d(CGCAAT**a**TAACGC)$^{3'}$	400	-5	0.026	-0.035

[a] λ_{ex} = 322 nm, [b] $\Delta\lambda_{em}$ = λ_{em} (dsDNA) - λ_{em} (ssDNA), [c] $\Delta\Phi$ = Φ(dsDNA) − Φ(ssDNA)

4.3 Oligonucleotide secondary structure

4.3.1 Melting temperatures

In order to assess the impact of the modified nucleobase on the stability of the synthesised oligonucleotides, their melting temperatures were measured and compared to the oligonucleotides containing the natural adenine base. Melting temperatures for the modified base with its correct base pair were measured. The base opposite the modified adenine was then changed to each of the other canonical bases to investigate the effect of base-pair mismatches on duplex stability.

The melting temperatures were recorded by heating a 1:1 mixture of the oligonucleotide and its complementary strand in a buffered solution (50 mM sodium phosphate pH 7.0, 0.1 M NaCl) from 20 °C to 90 °C at 1 °C/min, recording the absorbance at 260 nm. The same process was then repeated for annealing the duplex (90 °C to 20 °C). The data was differentiated, and the differentials fitted to Gaussian curves using OriginLab.

Table 4.4 shows the melting temperatures for the modified base in each of the sequences, with each canonical base opposite (T_m); it also shows the difference between the melting

curve containing the modified base with the natural oligonucleotide (ΔT_m)[‡]. This data is also shown graphically in Figure 4.6.

For each of the synthesised sequences, some duplex destabilisation was observed when the modified adenine was base-paired with thymine compared to the unmodified sequence (average destabilisation 6.24 °C). Although this is significant, it is less than the destabilisation associated with no base-pair being formed (average destabilisation 11.76 °C[§]).

[‡] Some of the melting temperatures of the unmodified duplexes were carried out by L. S. Griffin.
[§] Destabilisation calculated by averaging the differences between the base-paired (A-T) and non-base-paired (A-C, A-G, A-A) melting temperatures of the unmodified oligonucleotides.

Table 4.4 Melting temperatures of the modified oligonucleotides.

	Sequences	Opposite base (X)	T_m (control) /°C	T_m /°C	ΔT_m /°C
ODN1	$^{5'}$d(CGCTTG**a**GTTCGC)$^{3'}$	T	58.9	51.5	-7.5
		C	45.5	47.1	+1.6
1X	$^{3'}$d(GCGAAC**X**CAAGCG)$^{5'}$	A	47.7	46.0	-1.7
		G	47.5	42.8	-4.7
ODN2	$^{5'}$d(CGCTTC**a**CTTCGC)$^{3'}$	T	57.2	48.0	-9.2
		C	31.9	32.6	+0.7
2X	$^{3'}$d(GCGAAG**X**GAAGCG)$^{5'}$	A	46.8	39.9	-6.9
		G	53.0	47.1	-5.9
ODN3	$^{5'}$d(CGCAAG**a**GAACGC)$^{3'}$	T	55.7	49.5	-6.2
		C	42.7	44.4	+1.7
3X	$^{3'}$d(GCGTTC**X**CTTGCG)$^{5'}$	A	45.8	46.7	+0.9
		G	46.9	40.3	-6.9
ODN4	$^{5'}$d(CGCAAC**a**CAACGC)$^{3'}$	T	58.5	52.6	-5.9
		C	41.1	46.0	+5.0
4X	$^{3'}$d(GCGTTG**X**GTTGCG)$^{5'}$	A	48.7	40.7	-8.0
		G	50.7	47.0	-3.7
ODN5	$^{5'}$d(CGCAAT**a**TAACGC)$^{3'}$	T	50.2	47.7	-2.4
		C	35.1	31.9	-3.2
5X	$^{3'}$d(GCGTTA**X**ATTGCG)$^{5'}$	A	38.6	35.3	-3.3
		G	42.9	42.9	-0.1
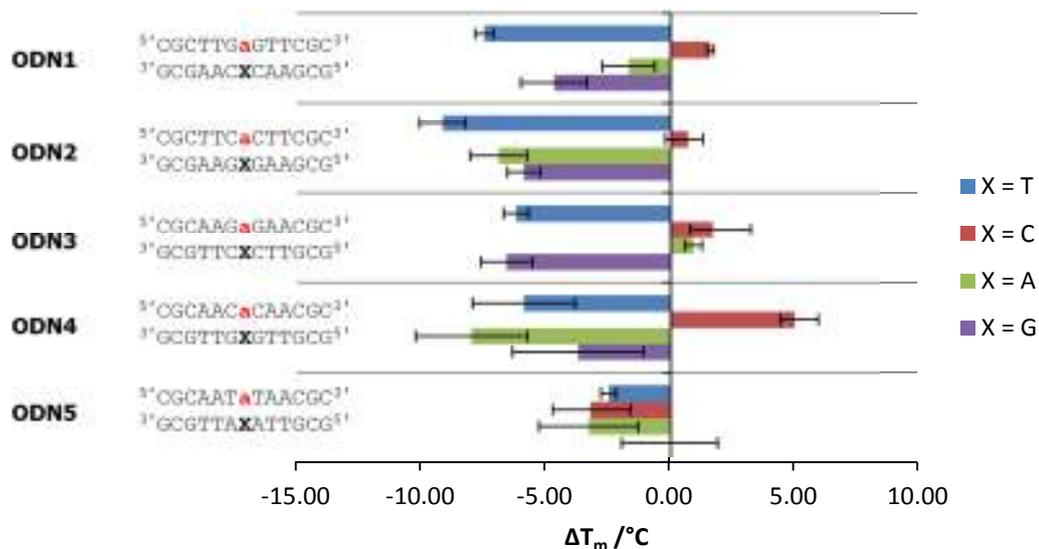
Figure 4.6 Differences in T_m between modified and unmodified oligonucleotides.

Error bars = $\pm\sqrt{\sigma(T_m1)^2 + \sigma(T_m2)^2}$, for X =T, T_m1 and T_m2: n = 6; for X = C, A or G, T_m1: n = 2, T_m2, n = 6.

It is also clear from Figure 4.6 that the base pair between the modified adenine and cytosine is significantly increased compared to the natural base (except in sequence **ODN5**). The increased stability may be indicative of some additional stabilising interaction between the modified adenine and cytosine. Looking at the difference between the modified adenine opposite T and the modified adenine opposite the other bases, however, shows no clear trend (Figure 4.7). The base-pairs with cytosine do not appear to be particularly stabilised compared with the other mismatches.
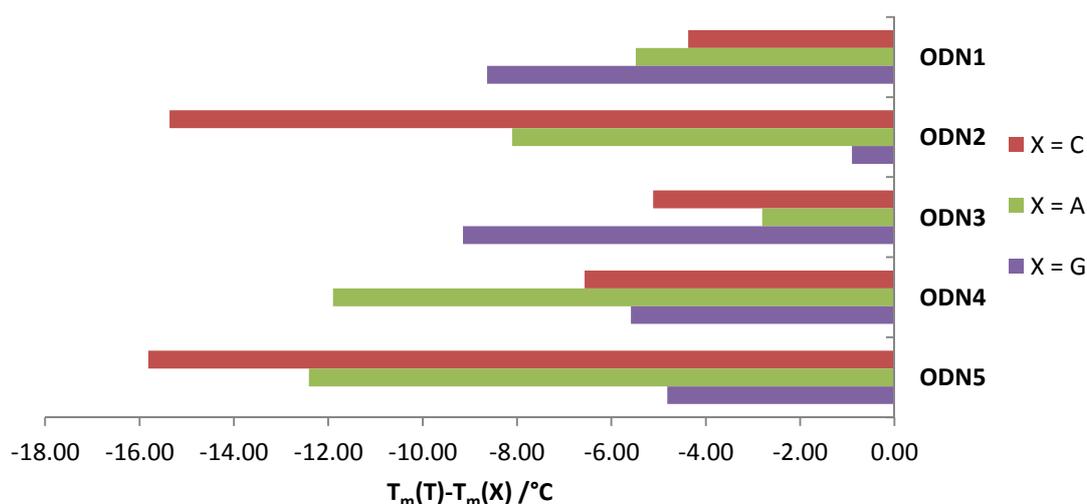
Figure 4.7 Difference between melting temperature of base-paired and mismatched base pairs in modified oligonucleotides.

100

4.3.2   Circular dichroism of the modified oligonucleotides

The first observation that is apparent upon comparison of the UV spectrum with the CD spectrum is that the absorbance peak for modified adenosine analogue is not present in the CD spectrum. This chromophore is planar and non-chiral, and chemically distant from the chiral centres of the molecule.

The single-strand fluorescent oligonucleotides exhibit very different CD spectra (Figure 4.8). Characteristically, B-form DNA displays a positive band at 260-280 nm and a negative band at 245 nm.[165] **ODN2**, **ODN4** and **ODN5** show CD spectra similar to that expected for B-form DNA. This suggests that the nucleotides have adopted the 2'-*endo* (S) conformation in solution, and it is likely that the bases are stacked in a single-strand helix (like in a B-DNA helix). Similar S-S stacked states have been previously observed for single-strand deoxyoligonucleotides which display B-form CD spectra.[171] Stacking interactions of single-strand oligonucleotides are thought to be favourable both enthalpically (due to the favourable van der Waals interactions) and entropically (due to the hydrophobic interactions).[172] **ODN3** has a strong positive peak at 280 nm, but only a weak negative peak at 245 nm. This is indicative of an A-type structure (or a mixture of A-type and B-type). It suggests that the nucleotides have adopted the 3'-endo (N) conformation, or a mixture of 2'-endo and 3'-endo conformations (S-N stacked). Oligonucleotides which form single-strand helices in solution are not as conformationally locked as those in double helices. In particular, the end nucleotides have greater conformational freedom and have been shown to adopt S-N stacked states at the terminal positions.[171] **ODN1** shows only an extremely weak CD spectrum, with a single peak at 280 nm. This suggests that **ODN1** does not have any specific secondary structure in solution.
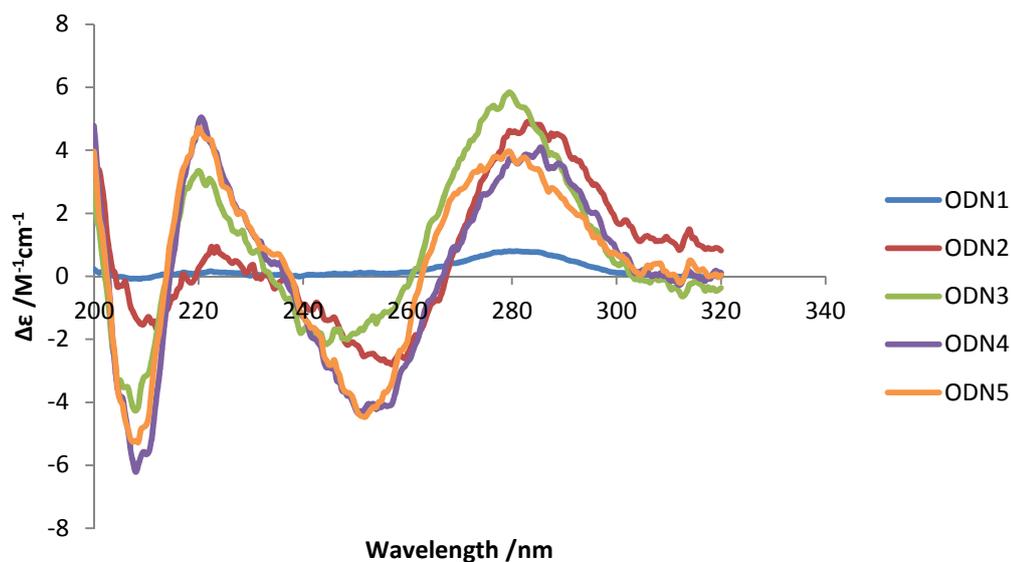
Figure 4.8 CD spectra of single-strand oligonucleotides (mean nucleotide values).

Circular dichroism spectra of the duplexes have a similar shape to the single-strand oligonucleotides, but with higher intensities. All the duplexes exhibit CD spectra consistent with B-type DNA structures. Most of the duplexes have CD spectra of comparable intensity; however the CD spectrum of **ODN2**-**2T** is much stronger than the other duplexes. Although its intensity is greater, it is very similar to that of the control duplex C2-2T. It may be that the two runs of four purines enhance base-stacking, giving a larger CD signal. Importantly, the CD spectra of the modified duplexes were closely matched to the unmodified duplexes. A representative comparison between the modified duplex and the control duplex is shown in Figure 4.10.
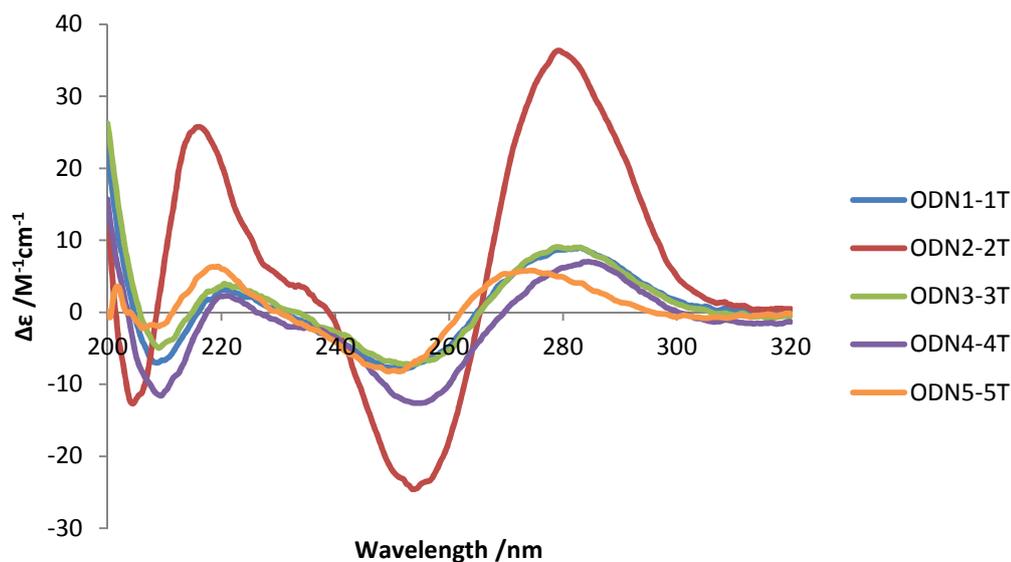
Figure 4.9 CD spectra of oligonucleotide duplexes (mean base-pair values) at 10 °C.
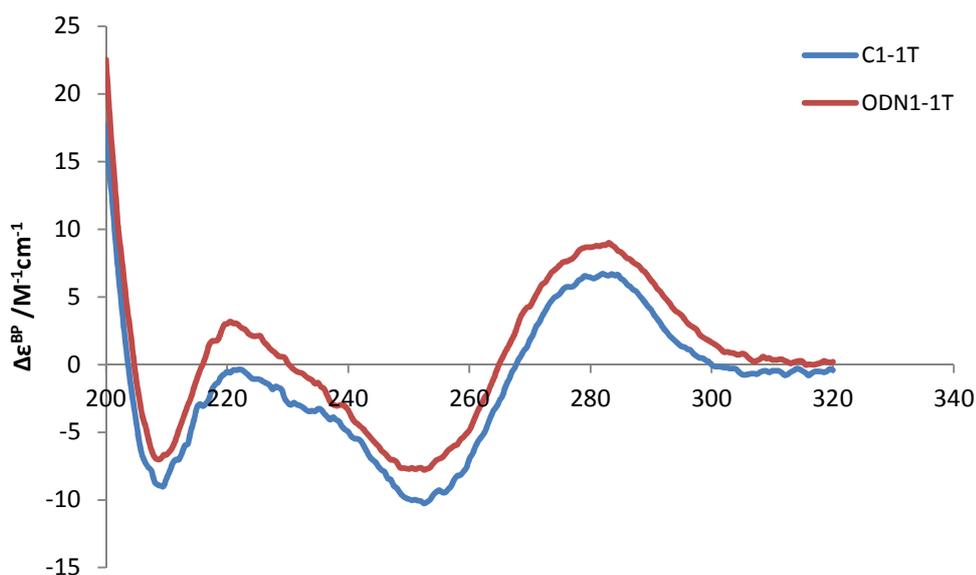
Figure 4.10 Circular dichroism spectra of **ODN1**-**1T** and **C1**-**1T** at 10 °C.

The circular dichroism spectra of the duplexes were measured at two temperatures, to give information for the annealed (10 °C) and melted (75 °C) oligonucleotides (Table 4.5). In all cases the duplexes displayed positive CD bands at *ca.* 285 nm and 220 nm, and negative bands at *ca.* 250 nm. The shape of the CD spectra generally stayed the same upon heating (Figure 4.11). This can be seen from the small changes in absorbance maxima and minima; however the intensity of the bands decreased significantly in all cases. This is consistent

with a decrease in secondary structure upon heating, although some secondary structure may still be present.
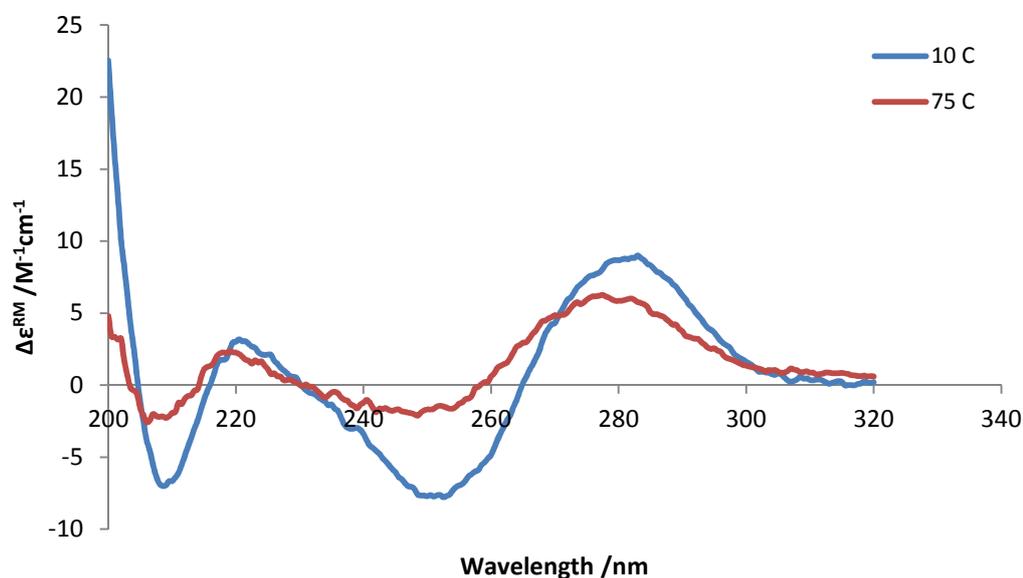
Figure 4.11 CD spectra of ODN1-1T at 10 °C and 75 °C.

Table 4.5 Major peaks in CD spectra of duplexes at 10 °C and 75 °C.

	Temperature	CD peaks					
	/°C	λ /nm	$\Delta\varepsilon$ /$M^{-1}cm^{-1}$	λ /nm	$\Delta\varepsilon$ /$M^{-1}cm^{-1}$	λ /nm	$\Delta\varepsilon$ /$M^{-1}cm^{-1}$
ODN1-1T	10	283	9.0	250	-7.7	221	3.1
	75	277	6.3	249	-2.1	219	2.4
ODN2-2T	10	279	36.3	254	-24.6	216	25.7
	75	279	12.6	253	-8.8	218	15.8
ODN3-3T	10	282	9.0	254	-7.2	221	3.9
	75	277	6.3	249	-2.0	218	2.2
ODN4-4T	10	285	7.0	255	-12.6	221	2.3
	75	282	4.9	250	-5.0	220	1.7
ODN5-5T	10	275	5.8	250	-8.2	219	6.3
	75	277	4.1	251	-5.3	220	2.4

4.3.3   Fluorescence quenching studies

4.3.3.1   Quenching studies on single-strand oligonucleotides

Potassium iodide, caesium chloride and acrylamide were all investigated as potential quenchers of the fluorescent nucleoside.
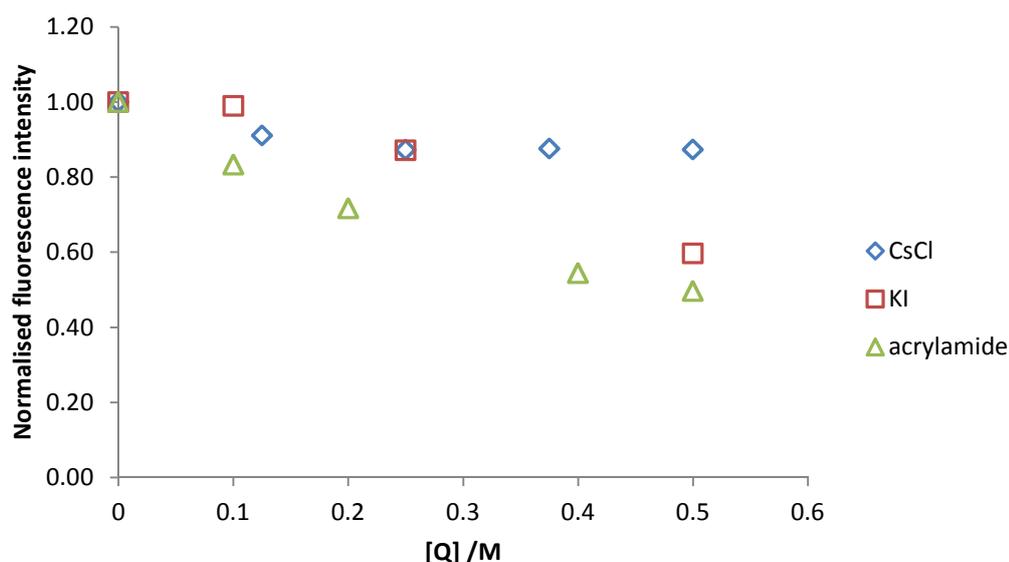
Figure 4.12 Graph showing changes in fluorescence intensity of **35** on addition of quenchers.

Caesium chloride was found to be the least efficient quencher, and acrylamide the most efficient for **35** (Figure 4.12). When 0.25 M CsCl was added, the fluorescence intensity decreased by ca. 15%, and did not change significantly if more CsCl was added after this point. Initially, addition of KI to **35** did not appear to give any quenching effect (at 0.1 M), but further additions showed some quenching effect. As caesium was found to be a poor quencher for the fluorescent nucleoside, and iodide was expected to be repelled by negatively-charged oligonucleotides, acrylamide was chosen as the quencher for further study.

Acrylamide solution was added portion-wise to the oligonucleotides and their duplexes in buffered solution at pH 7.0. There is a large variation in the quenching effect of acrylamide on the single-stranded oligonucleotides (Figure 4.13), which may be partially responsible for the large differences observed in fluorescence intensities. The plots of fluorescence

intensity against acrylamide concentration show broadly linear plots, although small variations due to static quenching are often difficult to observe if the static quenching constant is small. For a single fluorophore system, such as this, the degree of fluorescence quenching can be used as a guide to the exposure of the fluorophore to the solvent. From this it is clear that the different single stranded oligonucleotides have different exposures to solvent, suggesting perhaps that some secondary structure exists. CD spectroscopy indicates that all the oligonucleotides except **ODN1** have a helix-like secondary structure, which may shield the non-polar fluorophore from the aqueous solvent. The efficiency with which it can do this will be dependent on the nature of the secondary structure. The differences observed in the CD spectra are indicative of the range and type of secondary structures present. It is clear that the oligonucleotide with the most solvent exposure is **ODN3**, which is also the sequence with the highest quantum yield and fluorescence intensity. Additionally, **ODN3** has a strong CD signal, more like A-type structure than the other oligonucleotides, which may make the fluorophore more solvent exposed. **ODN1** shows no helical structure, which may mean that it is folded in such a way as to bury the non-polar fluorophore, meaning that it has little solvent exposure.
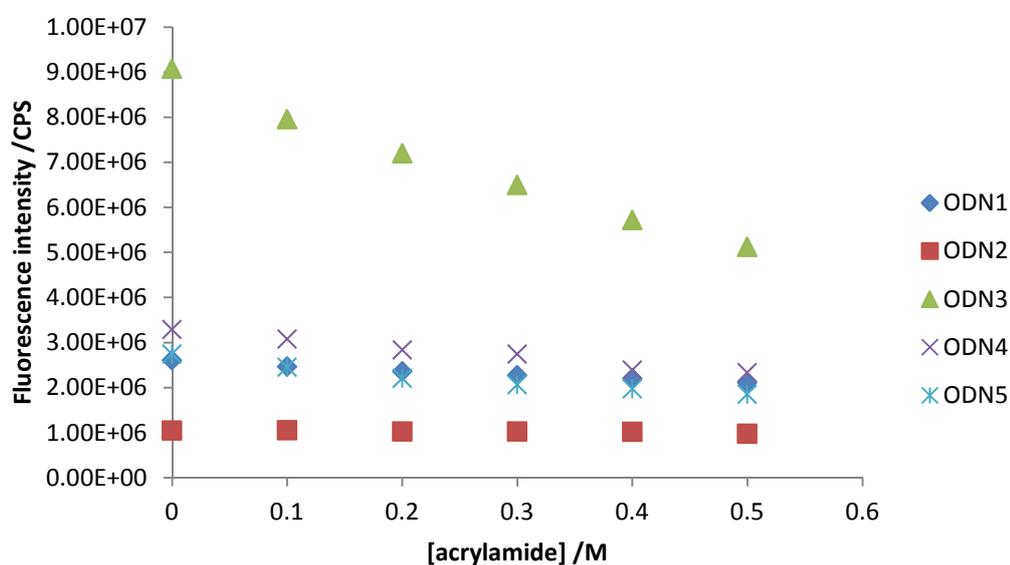
Figure 4.13 Quenching of single-strand oligonucleotides by acrylamide at λ_{400}.
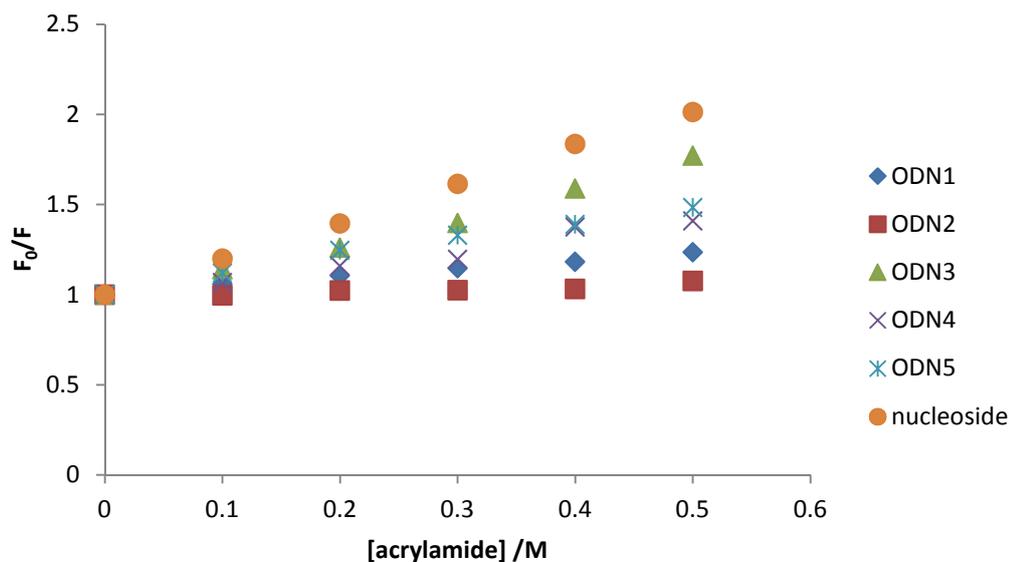
Figure 4.14 Stern-Volmer plots for single-strand oligonucleotides

There does appear to be some correlation between solvent accessibility and fluorescence intensity of the oligonucleotide, suggesting that this may be the origin of the variation in fluorescence observed for the single-strand oligonucleotides, rather than sequence dependent quenching (although the sequence may ultimately be responsible for the solvent exposure).

To test whether the oligonucleotides were forming some secondary structure, the dependence of fluorescence intensity on temperature was investigated. Fluorescence emission is also expected to vary with temperature as the number of collision events (which can cause non-radiative decay) increases with increased temperature. Therefore the effect of temperature on the fluorescence intensity of the unincorporated nucleoside was measured as a comparison to the oligonucleotides (Figure 4.15).
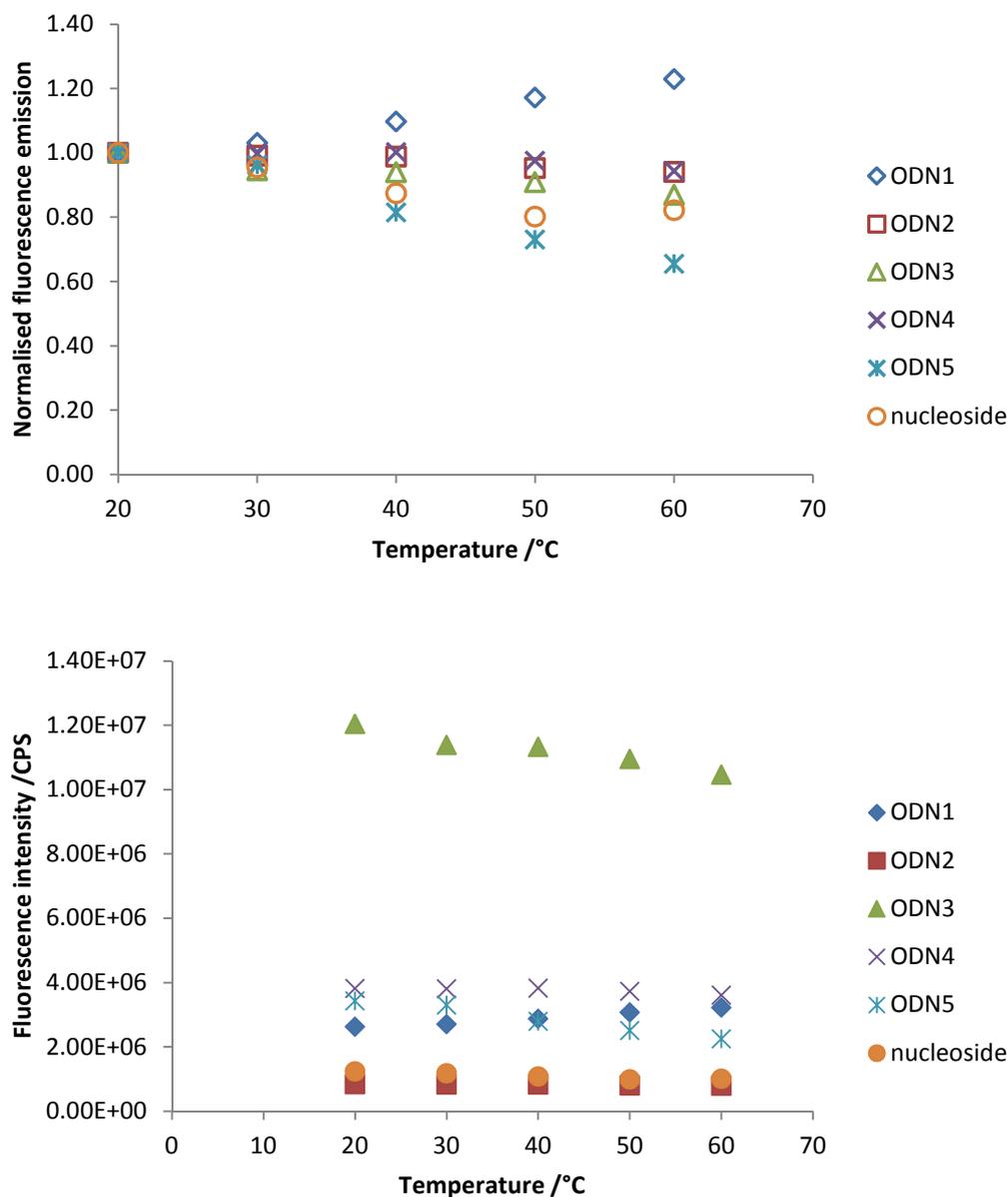
Figure 4.15 Dependence of fluorescence intensity on temperature: normalised (top panel) and actual (bottom panel) data given. Oligonucleotide fluorescence intensity measured at 400 nm, nucleoside fluorescence intensity measured at 365 nm.

From the normalised data it does appear that some "unfolding" occurs upon heating – particularly **ODN1** which shows a clear increase in fluorescence intensity upon heating. As **ODN3** has the highest fluorescence intensity and is the most quenched by acrylamide, it suggests that the more solvent-exposed fluorophores have higher emission intensities. Therefore, the increase in fluorescence intensity observed for **ODN1** upon heating suggests that the fluorophore may be more exposed to solvent at higher temperatures. As the curve is nearly straight (with a slight exponential curve), it suggests a non-cooperative

secondary structure unfolding. This is consistent with the CD spectrum which shows very little secondary structure, possibly causing the oligonucleotide to fold randomly to bury the fluorophore. However, alone this is clearly not enough to account for the differences in fluorescence intensity between the single strand oligonucleotides, and there is no correlation between the amount of quenching by acrylamide and the change in fluorescence intensity upon heating.

4.3.3.2 Quenching studies on oligonucleotide duplexes

In contrast to the single-strand oligonucleotides, the duplexes show little variation between their Stern-Volmer plots (Figure 4.16). This suggests that the differences between the fluorescence intensities of the duplexes are not because they have different solvent accessibilities. This supports the earlier suggestion that, once duplexed, the fluorescence intensities are dependent on the oligonucleotide sequence (specifically the neighbouring bases).
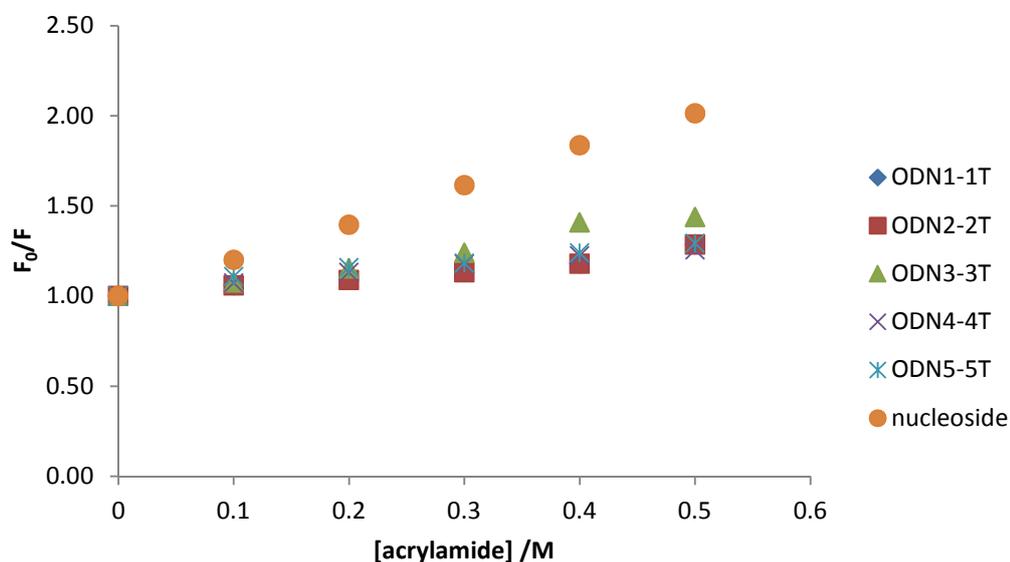
Figure 4.16 Stern-Volmer plots of modified oligonucleotide duplexes.

The effects of acrylamide quenching on the mismatched duplexes was not investigated. However, it is likely that in mismatched sequences (particularly the purine mismatches), the variation in fluorescence intensity is at least partially due to the solvent accessibility of the fluorophore.

4.4 Base-discriminating properties of the modified oligonucleotides

4.4.1 Fluorescence titrations

The changes in fluorescence spectroscopy on formation of a duplex were also examined. The complementary strand was titrated into a buffered solution of the modified oligonucleotide, and the fluorescence spectrum recorded after each addition. In each case, both the emission maximum and the fluorescence intensity change as the duplex is formed, and both level out at a 1:1 mixture of modified oligonucleotide and complementary strand. Since only small changes are observed after complete duplexation, this suggests that a stable 1:1 duplex is formed. Representative spectra are shown for **ODN1** with its complementary strand (**1T**) in Figure 4.17.

The changes in fluorescence intensity and emission maxima vary significantly between sequences, and depending on the base opposite the modified adenine (Figure 4.18). Because a wide range of sequences has been chosen to investigate these effects, trends have been identified in the fluorescence emission. However, definitively detecting trends would need further investigation.
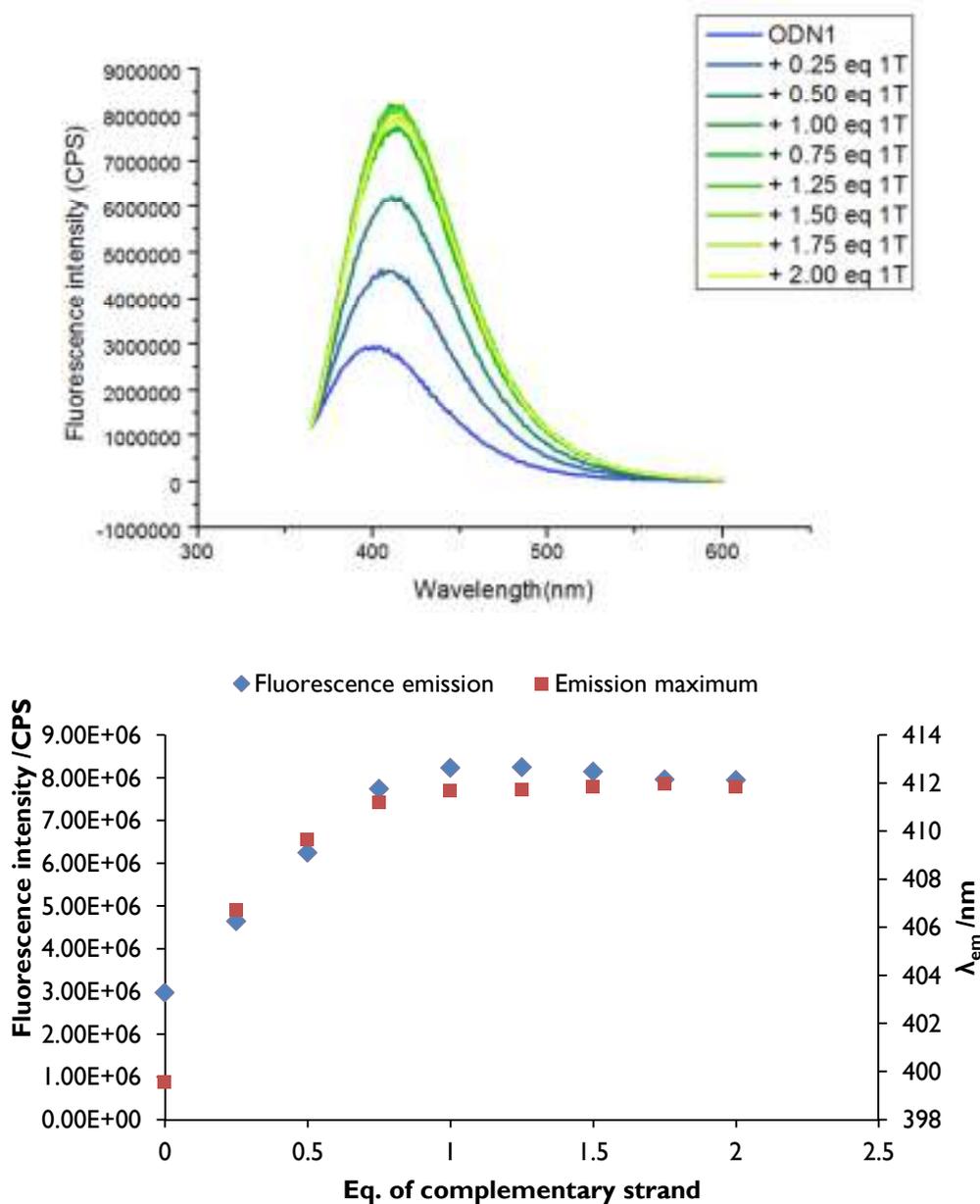
Figure 4.17 Top panel - Fluorescence emission spectra of **ODN1** with sequential additions of **1T**. Bottom panel - Changes in fluorescence intensity and emission maximum on addition of **1T** to **ODN1**.

When the emission intensities of the correct base-paired (**a**·T) duplexes is examined, it is clear that the fluorescence intensity is dependent on the bases adjacent to the fluorophore. **ODN1** and **ODN3** have sequences where the modified base is between guanines, and both **ODN1**-**1T** and **ODN3**-**3T** have similar emission intensities. Likewise, **ODN2**-**2T** and **ODN4**-**4T**, which both have the modified base adjacent to cytosine, have very similar (but lower) fluorescence intensities. **ODN5**-**5T**, with the modified adenine next to thymine, has a lower intensity still. This suggests that the pyrimidine bases have a

quenching effect on the fluorophore, and the two pyrimidine bases have different quenching effects. As this neighbouring base-dependent fluorescence is only seen with the correctly base-paired sequences, it suggests that the effect is dependent on the formation of the A-T base pair. For the base pair to be formed, the bases must be stacked in the duplex. This base-stacking effect, and the π-π stacking interactions associated with it, must be responsible for the sequence-specific fluorescence.

One trend that is clear from this data is that in each case, when the modified base is opposite cytosine, a quenching effect is observed. All the sequences show very low fluorescence (between 5×10^5 and 1.5×10^6) when a $c^7A \cdot C$ pair is present. The origins of this quenching effect will be discussed in subsequent sections.

The differences in fluorescence intensity of the single-strand oligonucleotides suggest that more complex factors affect the fluorescence intensity when the nucleotide is not in a duplex. The fluorescence intensity varies dramatically in the single stranded oligonucleotides, and this is reflected in the quantum yields (*cf*. Table 4.2). In particular, **ODN3** has much higher fluorescence than all of the other single-strand oligonucleotides. As both **ODN1** and **ODN3** have the modified nucleobase adjacent to guanine, it is clear that this is not the effect of neighbouring bases. The rest of the sequence in **ODN3** is shared with **ODN4**, suggesting that the large difference in fluorescence intensity between **ODN3** and the other oligonucleotides is either not dependent on the primary structure of the oligonucleotide, or it is some subtle effect from the combination of that specific sequence (or both). It therefore seems likely that some secondary structure (either helical or random) is responsible for at least some of the variation in intensity in the single strand oligonucleotides. The mismatched purine-purine base pairs also show significant difference between the sequences. Once again, there does not appear to be any clear pattern to these differences. In these cases, as the rest of the duplex is expected to be base-paired, it may be dependent on the specific conformation that the mismatched base pair adopts, and the solvent accessibility and local polarity of these conformations.
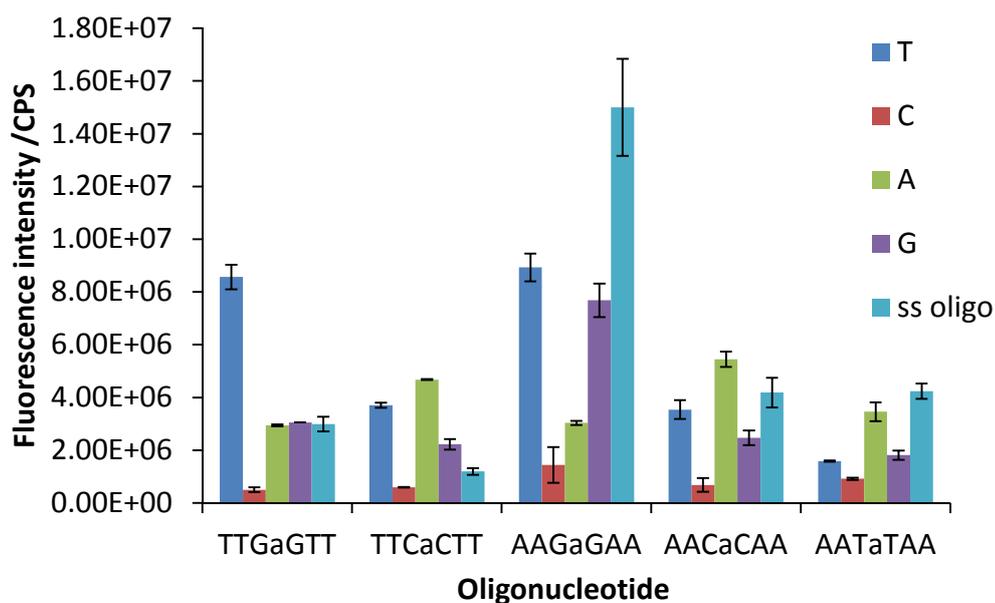
Figure 4.18 Fluorescence intensities of the duplexes of the modified oligonucleotides.

In order to better understand the variations in fluorescence emission of the single strand oligonucleotides which do not seem to be sequence specific, some further understanding of any non-specific secondary structure was required. Aside from sequence, many other factors can affect fluorescence intensity. The solvatochromic studies with the nucleoside show strong (if often non-specific) solvent dependence – although it was difficult to see patterns in the solvent dependence of the nucleoside, it is clear that the environment around the fluorophore has a pronounced effect on the fluorescent properties. Therefore, changes in fluorescence intensity may be indicative of changes in local polarity associated with duplex formation, and therefore be dependent on the position of the fluorophore relative to the duplex.

4.4.2 pH titrations

In order to understand the variation in fluorescence intensity between the matched and mismatched duplexes, the effect of pH on the fluorescence intensity was investigated. It has been previously observed by Seela and co-workers that 7-deazapurine-containing oligonucleotides show pH-dependent melting temperatures when mismatch base-pairs with cytosine are present.[173] Seela proposed that a wobble pair with cytosine could be formed if the 7-deazapurine is protonated (Figure 4.19). They had previously observed that the 7-deazaadenine heterocycle is considerably more basic than adenine (c^7dA pK_a = 5.3;

dA pK_a = 3.5). Using ^{15}N NMR studies, they determined that protonation occurs on the N1 position.[174] A·C mismatches can be formed with one hydrogen bond between the exocyclic amine of the adenosine and the N-3 position of the cytosine. Protonated adenosine-cytosine (A[+]·C) wobble pairs were proposed by Patel and co-workers for specific deoxyoligonucleotides, characterised using 1D[175] and 2D[176] NMR spectroscopy. The structure of the wobble pair wasn't clear until a crystal structure of a duplex containing an A·C pair was published by Brown and co-workers (Figure 4.20).[177] However, neither of these techniques gives definitive proof of protonation at N1. Generally, the protonated wobble pair forms only under acidic conditions, and is highly dependent on context within the sequence.[178] Wobble pairs are more often observed in RNA due to the wide range of RNA secondary structures, and indeed A[+]·C wobble pairs have been observed in RNA hairpin loops.[179] Protonation of adenosine is also required for the formation of some Hoogsteen base-pairs. The C·G*A[+] base-pair contains adenosine protonated at N1, and has been proposed for pH-dependent triplexes formed under acidic conditions.[180]
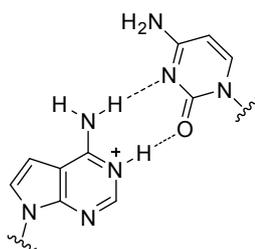
Figure 4.19 c[7]A[+]·C wobble pair proposed by Seela and co-workers.
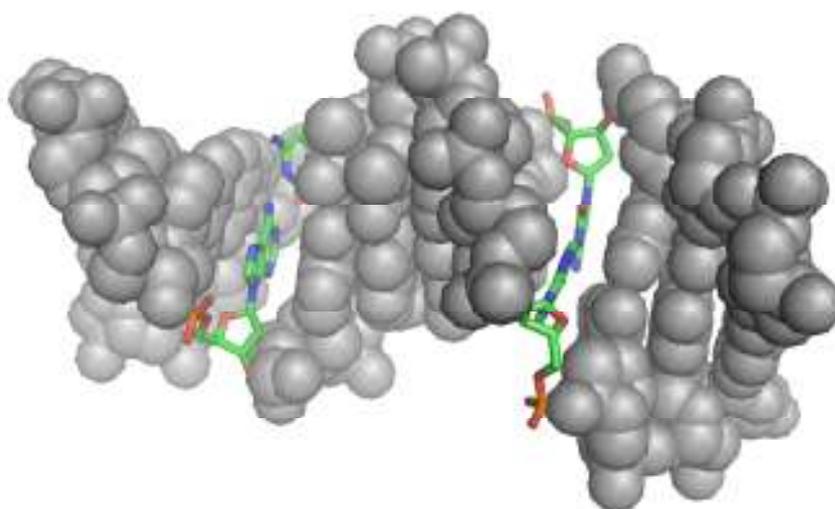
Figure 4.20 Crystal structure of the A[+]·C base-pair mismatches in the sequence d(CGCAAATTCGCG) reported by Brown and co-workers (PDB ID: 1D99).[177]

Investigation into the effect of pH variation on the nucleoside emission spectrum showed that protonation led to fluorescence quenching (*vide supra*). As base-pairs with cytosine also led to fluorescence quenching this suggests that a protonated $A^+ \cdot C$ wobble pair may be the origin of this quenching effect.

To investigate the effect of pH on the fluorescence of the oligonucleotides, the fluorescence emission spectra were measured in buffered solutions with varying pH. The spectra of ODN1 and its duplexes with T and C opposite the modified base are shown in Figure 4.21. From the data collated for all the modified oligonucleotides (Table 4.6), it is clear that irrespective of the sequence, the pK_as of the oligonucleotides are broadly the same. The single-strand oligonucleotides have pK_as around 5.3, significantly higher than the pK_a of the unincorporated nucleoside (*cf.* Figure 3.10) negatively charged phosphate backbone is known to increase the pK_a of nucleobases.[172] Upon formation of the correct base-pair with thymine, the pK_a decreases to *ca.* 4.6. This is expected, as protonation would destabilise the base-pair between the modified base and thymine. The base-pairs with cytosine show the opposite trend; the pK_a of the modified base increases to around 7.5. This means that the nucleotide is 100 times more basic when opposite cytosine. The increase in pK_a is consistent with stabilisation by formation of a base-pair. The 2006 Seela paper suggested that incorporation of the nucleobase into oligonucleotides will increase the pK_a of the heterocycle. However, these results suggest that the increase in pK_a is due to the formation of the base-pair; in effect, the base-pair acts as a template for the proton. It is also explains the origins of the base-discrimination between thymine and cytosine. As the fluorescence titration experiments were carried out at a pH close to the pK_a of the nucleotide in the A·C mismatch, approximately 50% of the base will be protonated. As it has been found that protonation causes fluorescence quenching, this gives rise to selective quenching of cytosine at pH 7.0. At pH 9.0, discrimination between the pyrimidine bases is sequence specific (Figure 4.22).
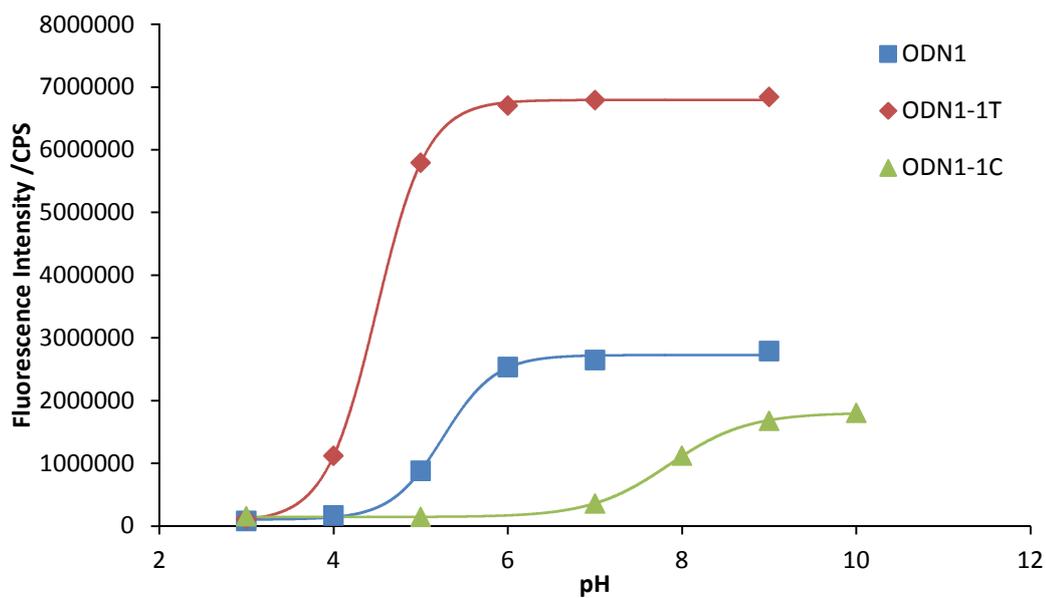
Figure 4.21 Effect of varying pH on the fluorescence emission spectra of **ODN1** and its duplexes at 400 nm.

Table 4.6 pKa values for the modified oligonucleotides and their duplexes, measured from their fluorescence spectra.

	Sequence	ODNX pK$_a$	ODNX-XT pK$_a$	ODNX-XC pK$_a$
ODN1	$^{5'}$d(CGCTTG**a**GTTCGC)$^{3'}$	5.25	4.49	7.85
ODN2	$^{5'}$d(CGCTTC**a**CTTCGC)$^{3'}$	5.59	4.62	7.71
ODN3	$^{5'}$d(CGCAAG**a**GAACGC)$^{3'}$	5.24	4.50	6.49
ODN4	$^{5'}$d(CGCAAC**a**CAACGC)$^{3'}$	5.18	4.41	7.88
ODN5	$^{5'}$d(CGCAAT**a**TAACGC)$^{3'}$	5.10	4.94	7.59
35	-	4.35	-	-
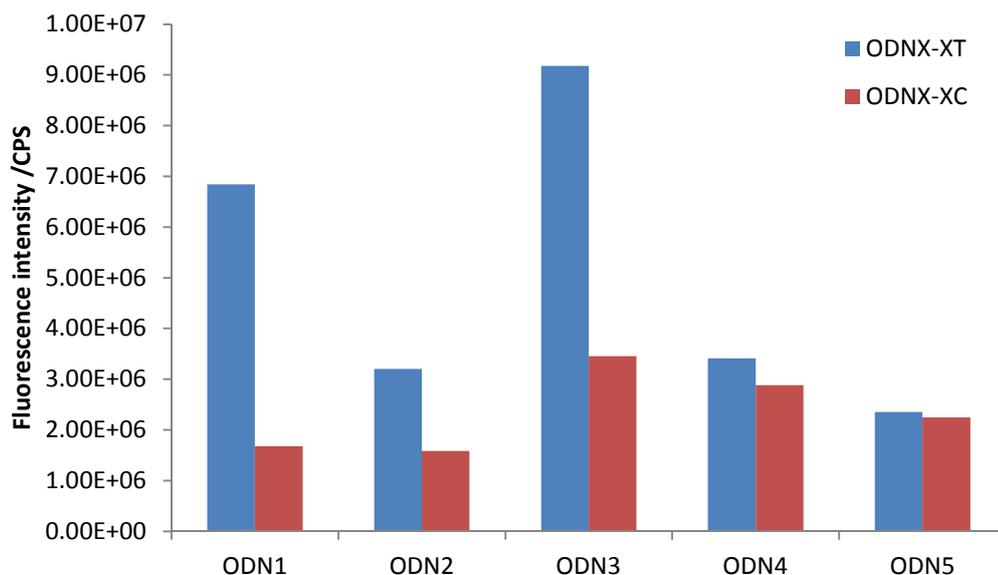
Figure 4.22 Fluorescence intensity of oligonucleotide duplexes at pH 9.0.

Consequently, this modified nucleotide (and potentially other related 7-deazaadenine systems) has useful base-discriminating properties over a relatively small pH range. This work has implications for many other reported 7-deazaadenine fluorescent probes, which may form mismatched base pairs below pH 7.0. This could give misleading results if used in an assay for SNP detection. Very few examples of base-discriminating nucleotides have been tested for their tolerance to variations in pH, so it is not known if they may produce inaccurate results if at basic or acidic pH. For example, base-discriminating 7-deazaadenosine nucleotides which rely on the formation or absence of a base pair causing conformational changes (such as the Saito pyrene-labelled nucleotides) may give a false positive at slightly acidic pH if and $c^7A^+ \cdot C$ base-pair is formed. Likewise, base-discriminating 7-deazaadenosine nucleotides like the ones described here, which rely on protonated $c^7A^+ \cdot C$ base-pairs causing fluorescence quenching, may give false negative results under slightly basic conditions.

4.5 Conclusions

UV/Vis and fluorescence spectroscopy of the fluorescence nucleotide in a variety of oligonucleotide 13mers showed that the fluorescence properties are dependent on the sequence context, whilst the UV/Vis spectra show little variation. The fluorescence

emission intensities of the single-strand oligonucleotides vary greatly, but with no clear relationship to the sequence. The fluorescence emission intensities of the duplexes, however, show a clearer correlation with the sequence. The duplexes containing the modified nucleotide flanked by guanosine (**ODN1** and **ODN3**) have the highest fluorescence emission intensity, followed by those duplexes where the fluorescent nucleotide flanked by cytidine (**ODN2** and **ODN4**). Comparing **ODN4** and **ODN5** it is clear that thymidine has a greater quenching effect than cytidine. The small differences between **ODN1** and **ODN3**, and **ODN2** and **ODN4**, which have the same flanking nucleotides but vary between purines and pyrimidines in the **a**±2 and **a**±3 positions, shows that the differences observed between sequences are primarily due to the adjacent nucleotides and not because of longer-range effects.

Melting temperatures of all the fully paired duplexes show significant destabilisation compared to the unmodified duplexes. The base-pair mismatches also show destabilisation, except in the case of cytidine mismatches, which mostly show an increase in stability compared to the control duplex.

Comparison of the circular dichroism spectra of the matched duplexes with the unmodified duplex spectra shows that they are extremely similar, with all duplexes displaying spectra typical of B-form DNA. CD spectra of the single-strand oligonucleotides reveal variations in secondary structure – this could be the cause of the differences observed between the fluorescence emission spectra. This is supported by fluorescence quenching experiments with acrylamide which show different solvent accessibility across the sequences, and temperature dependence of the fluorescence spectra where increasing the temperature can cause either an increase or decrease in fluorescence emission depending on the sequence. Notably, the circular dichroism spectrum of **ODN1** shows little secondary structure. **ODN1** is also only slightly quenched by acrylamide, and shows an increase in fluorescence intensity upon heating (consistent with greater solvent accessibly). This may indicate that **ODN1** is folded so as to bury the fluorophore, unfolding upon heating to increase its solvent accessibility.

Fluorescence emission of the mismatched base pairs with adenosine and guanosine were found not to be sequence dependent – there appears to be no clear pattern to the fluorescence intensities observed for the mismatched base-pairs. It is likely that subtle variation in conformation arising from base-pair mismatches changes the solvent-exposure of the nucleotide, and therefore the fluorescence properties. This could be further

investigated by fluorescence quenching studies on the mismatched duplexes. In every case, fluorescence emission arising from cytidine mismatches shows significant quenching.

The unusual results observed for the cytidine mismatches led to further investigation of their dependence on pH. It has been previously observed that adenosine is capable of forming $A^+ \cdot C$ mismatched pairs under specific conditions (generally acidic conditions). This has also been indicated for 7-deazaadenosines, which are significantly more basic than adenosines. We measured pK_a values for the single-strand oligonucleotides, their matched duplexes, and the $A \cdot C$ mismatches, and found that formation of $c^7 A^+ \cdot C$ mismatches increases the pK_a of the protonated species from *ca*. 5.3 (in the single-strand oligonucleotide) to *ca*. 7.5. The 7-deazaadenosine-cytidine mismatch acts as a template for protonation of the nucleobase. Protonation causes fluorescence quenching, giving rise to the observed base-discriminating properties.

5 Direct arylation of 2'-deoxyadenosine

5.1 Introduction

As previously described, direct arylation methodology is a useful tool for biaryl synthesis without the need for pre-functionalisation of one or both of the substrates (Section 1.4.3). It is particularly useful for heterocycles, where the heteroatoms often act as directing groups, helping to differentiate between C-H bonds. The regioselectivity can often be further enhanced by the addition of other metal salts, as described for the Pd/Cu-catalysed arylation of imidazole reported by Miura and co-workers. The addition of excess CuI biases reactivity towards arylation at C2, rather than C5 of 1-methylimidazole.[87] The amount of CuI required, however, is dependent on the type of heterocycle (Scheme 5.1).
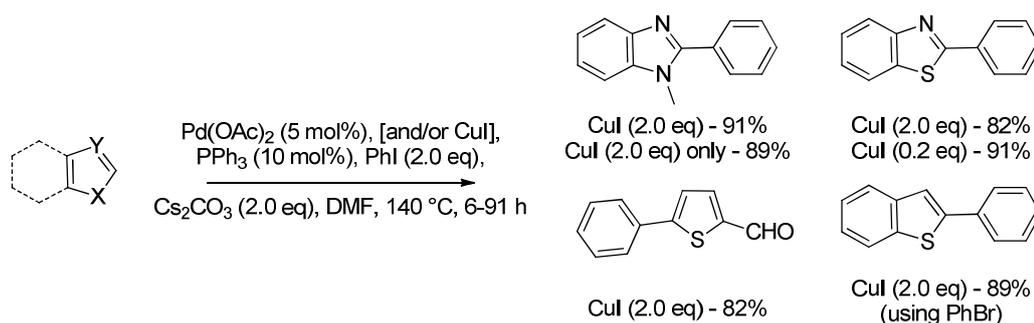
Scheme 5.1 Requirement for CuI additives in the direct arylation of a variety of heterocycles.

This raises questions about the role of CuI in the direct arylation reactions. The mechanism proposed for the direct arylation of adenosine reported by our group (and that of others) does not adequately account for the requirement of excess CuI (Scheme 5.2). However, it is clear that CuI can form relatively stable Cu^I complexes with 2'-deoxyadenosine, therefore excess CuI is needed for forming any organocuprate species. It was found that 2'-deoxyadenosine forms relatively stable complexes with Cu^I, coordinated through the exocyclic amine and N7 positions. When the complex [Cu(*N6,N7*-dA)$_2$]I is subjected to the reaction conditions (Pd(OAc)$_2$, PhI, Cs$_2$CO$_3$, piperidine, DMF, 80 °C), only a trace of product is formed. However, when the complex [Cu(*N6,N7*-dA)$_2$]I.CuI (with Cu:dA = 1:1) is subjected to the same conditions, the 8-arylated product is formed in 41% yield. The additional CuI is thought to be destabilising the complex, forming monoligated (and

therefore more reactive) Cu species. Sonogashira reactions with 8-bromoadenosine are known to proceed with very low loadings of CuI (2 mol%), suggesting that coordination to CuI is not necessarily problematic, although the reaction conditions are different in this case (other competing ligands such as alkynes and amines are present).
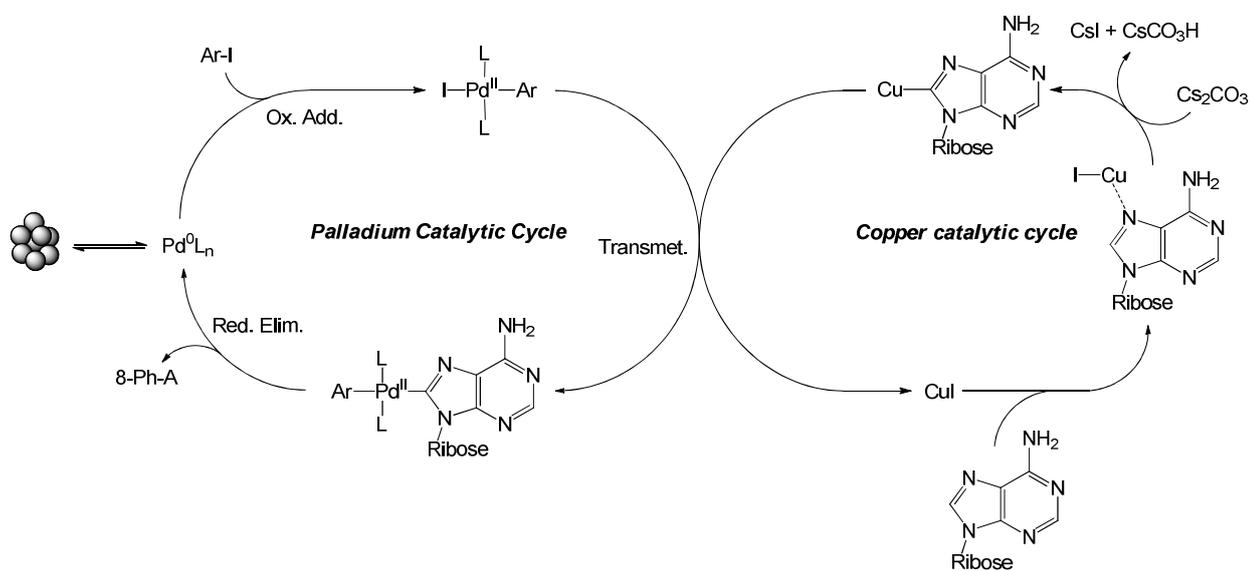
Scheme 5.2 Proposed mechanism for the direct arylation of adenosine.

Bellina, Rossi and co-workers proposed that the requirement for excess CuI in these reactions was due to reversible cupration of the heterocycle. However, it seems unlikely in the presence of excess base that the deprotonation reaction will be reversible.

More recently, a system which employs both palladium and copper in sub-stoichiometric amounts has been reported by Huang and co-workers.[181] Their system uses PXPd (0.25 mol%) and Cu(Xantphos)I (1 mol%), and was successful for the arylation of a variety of heterocycles affording products in good to excellent yields (Scheme 5.3). However, these conditions were not found to be suitable for the arylation of nucleosides.

PXPd (0.25 mol%), Cu(Xantphos)I (1 mol%),
ArX (1.5 eq), Cs$_2$CO$_3$ (2.5 eq)

Tol, 100 °C, 6^{-22} h

PXPd = PdCl$_2$(^{t}Bu$_2$PCl)$_2$

X = O, S, NH
20 examples, 48–99%
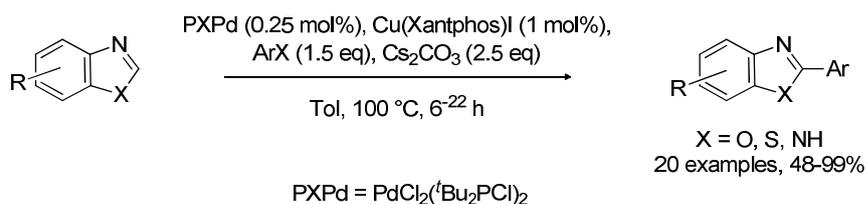
Scheme 5.3 Co-catalytic Pd/Cu arylation of heterocycles.

It seems that the nature of the copper source, as well as the identity of the substrate, may dictate whether excess copper salts are required. However this area is not fully understood, and more work must be done to determine the role of CuI in this reaction.

One of the great advantages of the direct arylation methodology is that it potentially allows for the functionalisation of biological molecules, specifically biopolymers, without the need for incorporation of a modified monomer. In order to apply the chemistry to more complex biological molecules, milder conditions must be developed because of the temperature-sensitive nature of biomolecules. Understanding the reaction mechanism is key to development of improved, milder reaction conditions.

5.2   Low temperature direct arylation of 2'-deoxyadenosine

A new major challenge in this area is the direct arylation of unmodified oligonucleotides. A major advantage of the methodology developed within our group is the relatively mild conditions employed; the direct arylation of 2'-deoxyadenosine can be carried out at 80 °C. This is in stark contrast to many analogous arylation procedures, which generally employ temperatures between 100 and 160 °C. However, to be applied to nucleotides or oligonucleotides the conditions must be even milder, at most 50-60 °C. To this end, some optimisation of the reaction at lower temperatures has been carried out and is reported here.

Work carried out by T. E. Storr (York) indicated that the role of piperidine in the reaction at 80 °C was to reduce the Pd^{II} pre-catalyst to the Pd^0 active catalyst as previously observed by Buchwald and others.[182] Therefore, the use of Pd^0 catalysts in this reaction was investigated (Scheme 5.4). Initial results showed excellent conversion after 24 h at 60 °C using Pd$_2$(dba-4-OMe,4'-CF$_3$)$_3$, a catalyst developed in our group (unpublished),[183] however

122

these results were not found to be repeatable. It is thought that the catalyst was contaminated with Pd^0 nanoparticles, and that these may be positively influencing the reaction. Results from a screen of Pd^0 catalysts are shown in Table 5.1.

A variety of Pd^0 catalysts (Table 5.1, entries 2-4) showed similar activity. The increased yield when $Pd(PPh_3)_4$ was used as the catalyst under air (phosphine oxidation leads to release of Pd^0 atoms which aggregate via Ostwald ripening processes), the irreproducible yield (entry 1) and our suspicions about nanoparticle contamination with this catalyst also led to Pd^0 nanoparticles being tested in the reaction (entries 6-8). The low activity observed with Pd^0 nanoparticles stabilised by polystyrene is thought to be partially due to the low solubility of these nanoparticles in DMF. Pd^0 nanoparticles stabilised by PVP (poly-N-vinyl-pyrrolidone) have been shown within our group to act as an effective catalyst source for heterogeneous Suzuki coupling.[184] The particles can be reliably synthesised from $H_2[PdCl_4]$ with small particle size distributions (Figure 5.2). The Pd^0-PVP particles were found to give a moderate yield in the direct arylation of 2'-deoxyadenosine at 60 °C (entry 8). This catalyst had previously been tested as a palladium source in the direct arylation of adenosine at 120 °C. It was found that much lower catalyst loadings could be used in this reaction when Pd^0-PVP was the catalyst source. However, when the direct arylation at 60 °C was carried out with this palladium source at 0.7 mol%, a significant drop off in yield was observed (entry 7). These conditions are not fully optimised; therefore an ideal catalyst loading may lie between these two values. As catalysis is thought to occur on defect positions on the nanoparticle surface, the number of active Pd centres is significantly lower than the catalyst loading suggests.
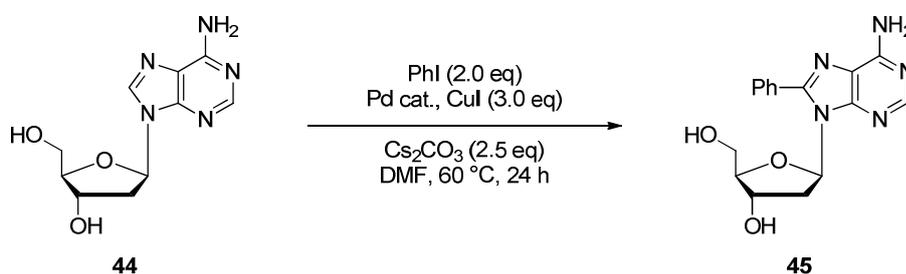
Scheme 5.4 Screening reaction for Pd catalysts for the low temperature direct arylation of 2'-deoxyadenosine.

Table 5.1 Screening of palladium catalysts in direct arylation of 2'-deoxyadenosine at 60 °C.

	Palladium catalyst	Loading /mol%	Yield /%
1	Pd$_2$(dba-4-OMe,4'-CF$_3$)$_3$ (original sample)[a]	5	81
2	Pd$_2$(dba-4-OMe,4'-CF$_3$)$_3$ (pure by microanalysis)	5	24
3	Pd$_2$(dba-4-,4'-OMe)$_3$	5	20
4	Pd(PPh$_3$)$_4$	5	19
5	Pd(PPh$_3$)$_4$[b]	5	30
6	Pd^0/polystyrene nanoparticles	5	23
7	Pd^0/PVP nanoparticles	0.7 (11 wt% Pd)	20
8	Pd^0/PVP nanoparticles	5 (11 wt% Pd)	52

[a] Complex likely contaminated by Pd nanoparticles (formed in the synthesis of the complex or by degradation); [b] Reaction carried out under air
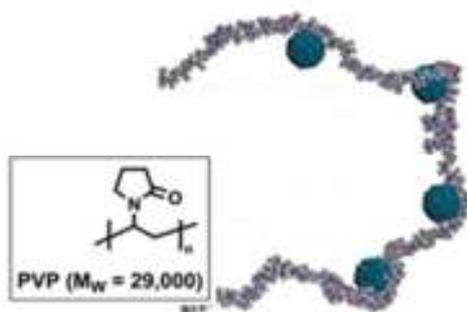
Figure 5.1 Representation of Pd nanoparticles supported by PVP.

Modelled using computational approaches by Prof. A. Lee (Cardiff)
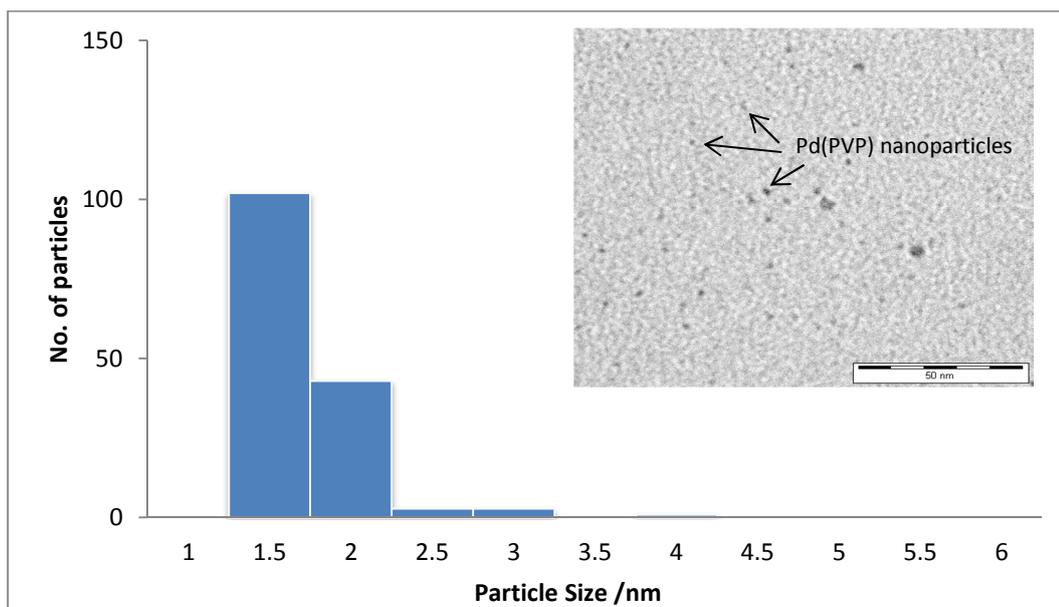
Figure 5.2 Histogram showing size-distribution of synthesised Pd(PVP) nanoparticles (measured using the PSA macro for ImageJ, minimum particle size set to 1.5 nm). Inset – TEM image of the Pd(PVP) nanoparticles.

A small variety of phosphine ligands, including both electron-withdrawing and electron-donating phosphines were also tested in the reaction, using $Pd_2(dba)_3$ as the palladium pre-catalyst and a 2:1 ratio of phosphine:Pd (Table 5.2). In all cases, a decrease in yield was observed when a phosphine was added to the reaction mixture. This is another indicator of Pd nanoparticle involvement. Coordination of the Pd catalyst by phosphine can slow the formation of Pd nanoparticles by stabilising the mononuclear species (especially under rigorously inert reaction conditions).

Table 5.2 Screening of phosphine ligands in direct arylation of 2'-deoxyadenosine at 60 °C

Phosphine	Yield /%
XPhos	31
$P(C_6F_5)_3$	9
PCy_3	7
none	42

$Pd_2(dba)_3$.dba (2.5 mol%), phosphine (10 mol%), other conditions as Scheme 5.4.

To investigate whether Pd nanoparticles could be the active species in the direct arylation reaction at 80 °C with Pd(OAc)$_2$ and piperidine, aliquots were removed from the reaction mixture and added to PVP, (M$_w$ = 29000), before the solvent was removed. Addition of PVP stabilises any nanoparticles in solution, so particle growth does not occur just through removal of the solvent. TEM images of the samples showed small particles, *ca.* 2 nm. Higher resolution TEM would be necessary to fully assess the size and shape of the particles observed, but the size distribution is very similar to that observed for the pre-synthesised Pd(PVP) particles (Figure 5.2). The group had previously observed Pd nanoparticles in the direct arylation reaction of adenosine at 120 °C with Pd(OAc)$_2$.[83]
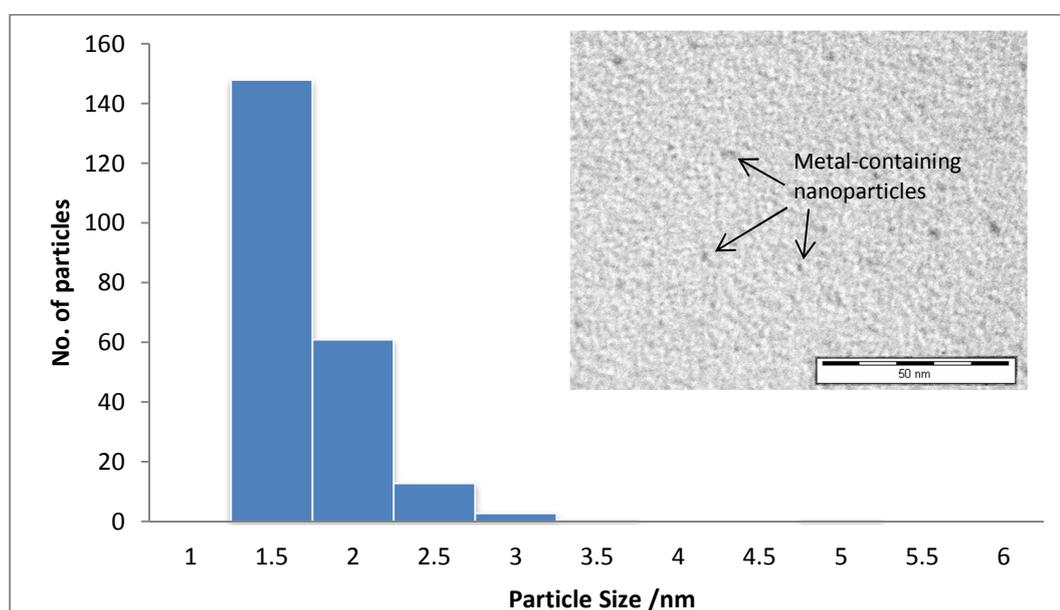
Figure 5.3 TEM images Histogram showing size-distribution of samples taken from Pd(OAc)$_2$ with piperidine reaction. Lower image shows an expanded region metal-containing nanoparticles (measured using the PSA macro for ImageJ, minimum particle size set to 1.5 nm). Inset – TEM image of the metal-containing nanoparticles.

A small library of aryl iodides was tested in this reaction at lower temperature to investigate the substrate scope at this temperature with the Pd^0-PVP catalyst (Table 5.3, Scheme 5.5). The reactions were carried out over 48 h to maximise the conversion however this did not lead to any enhancement in the product yields.

Table 5.3 Investigation of substrate scope at 60 °C over 48 h with Pd^0(PVP).

Ar-	Yield /%
Ph-	50
4-MeO-C$_6$H$_4$-	55
3-CF$_3$-C$_6$H$_4$-	25*
4-F-C$_6$H$_4$-	53
4-Tol-	38

*contaminated with *N*-arylated product.
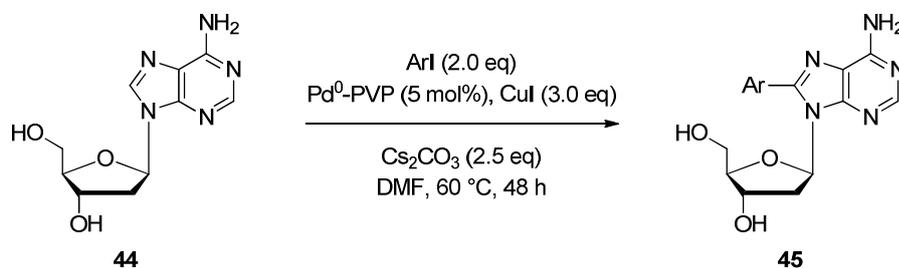
Scheme 5.5 Investigation of substrate scope in direct arylation of 2′-deoxyadenosine at 60 °C.

The significant variation in the yield, dependent on the aryl iodide used, suggests the dependency of the conversion on the palladium catalytic cycle. The rough trend seen here is in contrast to the yields observed for the direct arylation at 80 °C with Pd(OAc)$_2$ and piperidine, which gave the *m*-CF$_3$ product in 99% isolated yield.[86] It may be that cupration or transmetallation steps are slower at 60 °C, meaning that the Pd complexes resulting from rapid oxidative addition of aryl iodides may be present in solution longer, increasing the likelihood of homocoupling or protonation (Scheme 5.6). It is important that the rates of both catalytic cycles are closely matched in order to achieve the optimum yield.
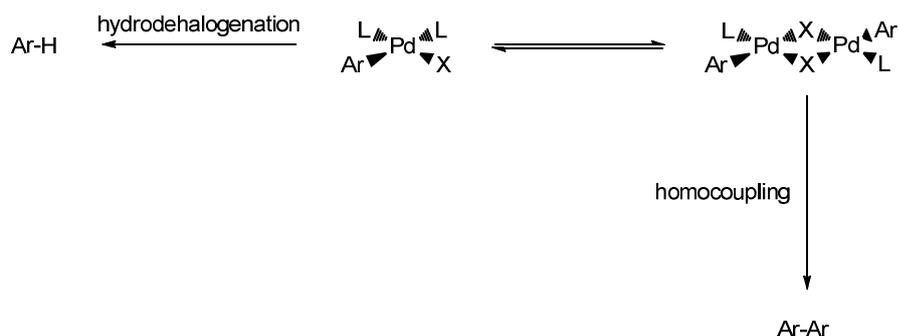
Scheme 5.6 Some deactivation pathways of Pd oxidative addition products.

5.3   Role of Cu[I] salts

As alluded to previously, the role of the stoichiometric CuI additive in these reactions is not fully understood.   Understanding the reaction mechanism and identifying the rate-determining step is crucial to the development of milder conditions.   Changing the temperature might also cause a change in the rate-determining step.   Deuteration studies were carried out in order to determine the efficiency of the deprotonation step at lower temperatures (Scheme 5.7 and Table 5.4).
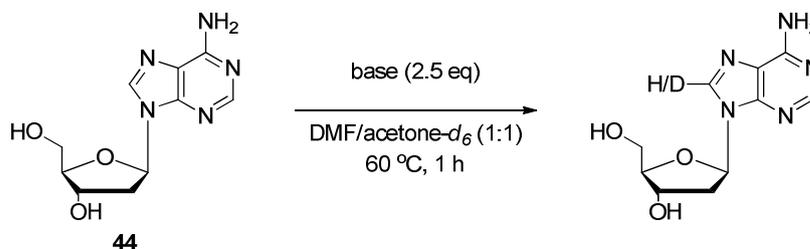
Scheme 5.7 Deuteration of 2'-deoxyadenosine.

Table 5.4 Deuteration of 2'-deoxyadenosine at 60 °C.

Base	% Deuteration
Cs_2CO_3 (80 °C)	96[86]
Cs_2CO_3	75
Cs_2CO_3 (extrafine)	84
Rb_2CO_3	19

Rb$_2$CO$_3$ was considered as an alternative to Cs$_2$CO$_3$, as the change in rate of transmetallation may increase the overall rate of reaction.[185] 2'-Deoxyadenosine was deuterated much less efficiently with rubidium carbonate than with caesium carbonate, as expected. Although it still seemed viable as a base, albeit much weaker, when it was used in the direct arylation reaction under the standard conditions, no product formation was seen. It is expected that coordination of Cu^I to the N7 position of the nucleoside will lower the pK$_a$ of the C8 proton, and aid deprotonation. However, attempts to test this using Rb$_2$CO$_3$ and CuI failed, as coordination of copper to the substrate lead to broad NMR signals which could not be integrated. Substitution of Cs$_2$CO$_3$ with Rb$_2$CO$_3$ in the direct arylation reaction led to oxidation of the CuI to a Cu^{II} species, and only a trace formation of product.

It was proposed that the need for excess CuI could be due to a reaction of CuI with Cs$_2$CO$_3$ to form Cu$_2$CO$_3$, an unreported and presumably highly reactive species (Scheme 5.8). The copper carbonate formed could be responsible for an intramolecular deprotonation-cupration of the adenosine heterocycle. If the equilibrium for formation of this copper carbonate intermediate lies towards copper iodide then high concentrations of CuI would be required for small amounts of copper carbonates to be formed.

$$Cs_2CO_3 \quad + \quad 2CuI \quad \underset{\xrightarrow{\hspace{2cm}}}{\overset{DMF}{\rightleftharpoons}} \quad Cu_2CO_3 \quad + \quad 2CsI$$

Scheme 5.8 Proposed reaction of Cs$_2$CO$_3$ with CuI.

To test this hypothesis, the reaction of Cs$_2$CO$_3$ with CuI was examined by *in situ* IR spectroscopic analysis ("ReactIR"). When Cs$_2$CO$_3$ was mixed with CuI at the same concentrations as in the direct arylation reaction, as expected a new IR band, corresponding to a metal carbonate was observed at 1652 cm^{-1} (Figure 5.5 and Figure 5.6). The solution also changed colour from colourless (with the Cs$_2$CO$_3$ as a white precipitate) to dark yellow (also with precipitate present) (Figure 5.7). No IR band is observed for the Cs$_2$CO$_3$ alone, probably due to its extremely low solubility in DMF. This seemed to confirm the theory of formation of a soluble copper carbonate species. However, a second new IR band was also observed upon addition of the CuI. The second IR band appeared at 2350 cm^{-1}, which corresponds directly to carbon dioxide. Both IR bands quickly level-off, possibly due to saturation of the solution. This suggests that the new carbonate species may be decomposing to CO$_2$, which could be a pathway leading to formation of a possibly inactive copper species leading to the requirement for excess CuI (Scheme 5.9). When CuI was

replaced with Cu_2O in the direct arylation reaction, no product was formed. This provides evidence for a deactivation pathway. However, it has not been possible to prove if Cu_2CO_3 is the active copper species, or merely a step on the deactivation pathway, as it is a previously unknown and highly reactive complex. It is possible that the cupration mechanism could be similar to the AMLA-6 (Ambiphilic Metal Ligand Activation) and CMD (concerted metalation-deprotonation) transition states proposed by Davies and Macgregor[186] and Fagnou[187] (Figure 5.5). Despite the unanswered questions, it seems highly likely that this is a significant step within the overall context of the arylation chemistry.

Initial coordination mode　　　　**AMLA-6 or CMD intermediate**
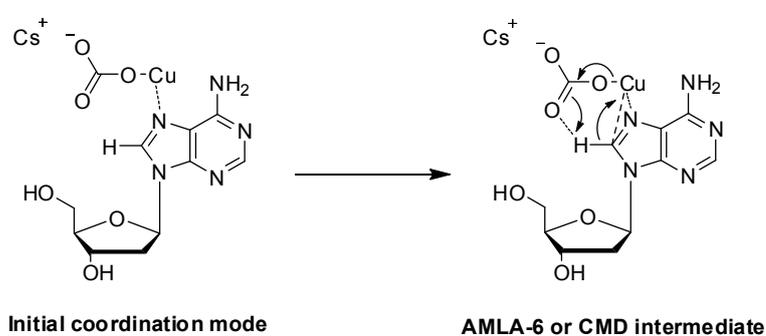
Figure 5.4 Proposed concerted deprotonation-cupration by Cu_2CO_3 (in DMF).
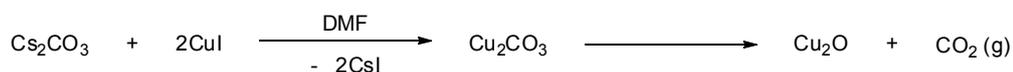
Scheme 5.9 Evolution of CO_2 (g) from CuI and Cs_2CO_3.
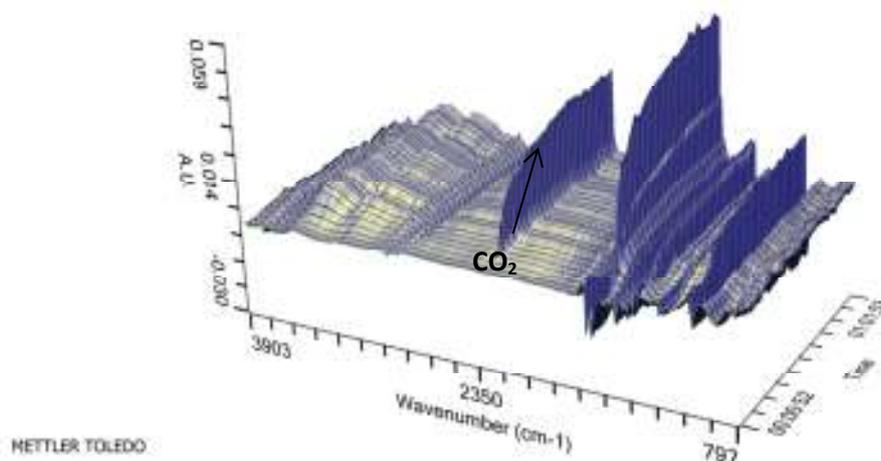
Figure 5.5 ReactIR spectra following the reaction of CuI with Cs_2CO_3 in DMF over time at rt. The addition of CuI was made at 05:50 mins.
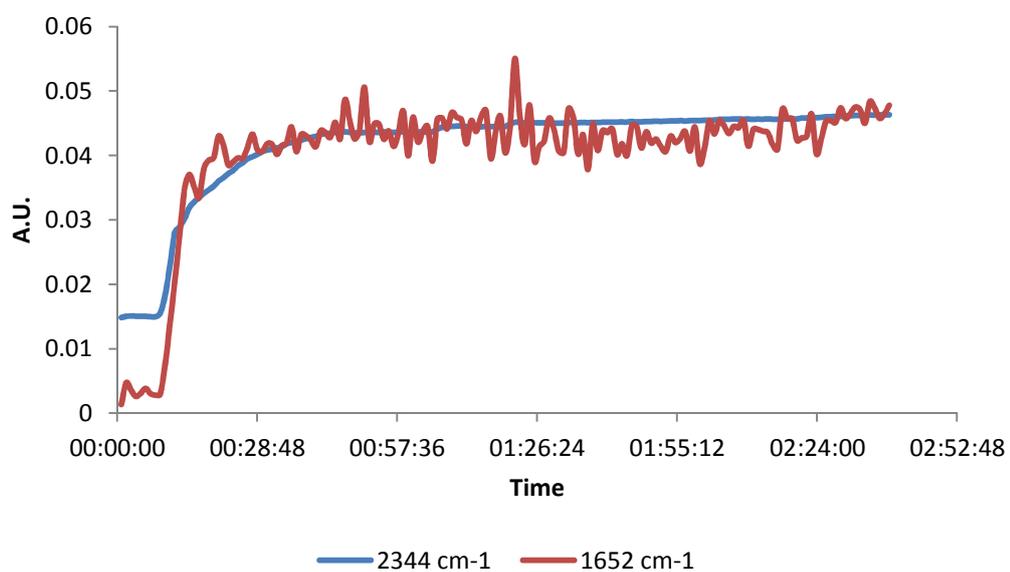
Figure 5.6 Change over time of selected IR bands. The addition of CuI was made at 05:50 mins.
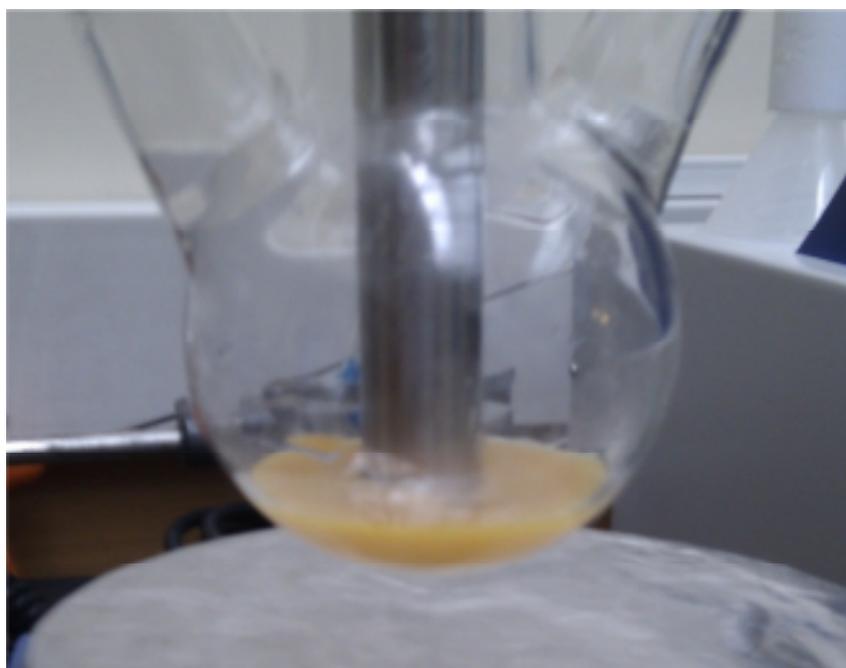
Figure 5.7 Reaction mixture containing CuI and Cs_2CO_3 in DMF (the reaction monitored by ReactIR).

5.4 Conclusions

It has been shown that the direct arylation of 2′-deoxyadenosine can be successfully carried out using Pd nanoparticles supported by PVP at 60 °C. The yields are moderate, however direct arylation chemistry at this temperature is rare, so even moderate yields are useful, particularly in more structurally complex systems that could be tested in the future (*e.g.* oligonucleotides). Evidence suggests that Pd nanoparticles play an important role in the catalysis, although it is not clear whether the catalysis takes place on the surface of the particle, or if the nanoparticles act as a reservoir of mononuclear Pd - this requires a comprehensive study in its own right, akin to that conducted for Suzuki-Miyaura cross-coupling catalysed by Pd(PVP).[184] Some insight has also been gained into the role of excess CuI. CuI was found to react with Cs_2CO_3 to form a species consistent with a metal carbonate, as well as CO_2 (g). This is a deactivation pathway, which forms the inactive copper salt Cu_2O. The formation of Cu_2CO_3 could be vital to the success of the reaction. DFT studies on the copper-catalysed arylation of related heterocycles such as benzoxazoles suugests that the reaction proceeds *via* a four-membered transition state, and the deprotonation is mediated by an alcohol.[188] Protection of the nucleoside sugar would indicate whether the sugar hydroxyls are acting as the alcohol mediating this reaction.

6 Conclusions and Future Work

An overall summary of the thesis is given below, divided into four sections, with a summary and perspective on future directions for the research in the final section.

6.1 Design and synthesis of novel fluorescent 7-deazaadenosine nucleosides

A novel class of 7-modified-7-deazaadenosine nucleosides containing a diarylacetylene fluorophore at the 7-position was proposed (Figure 6.1). The fluorescent nucleosides were designed to be sensitive to their local chemical environment, specifically to DNA hybridisation, thus allowing the detection of base-pair formation. Such fluorescent nucleosides are suitable for identification of single-nucleotide polymorphisms (SNPs).
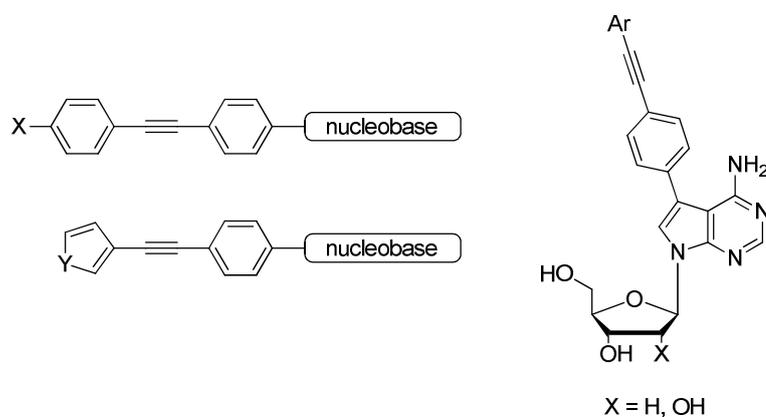
Figure 6.1 Novel Rigid-rod fluorescent labels for 7-deazapurine nucleosides.

The 7-deazaadenosine nucleosides were synthesised from the corresponding 7-iodo-7-deazaadenine nucleoside by Pd-catalysed cross-coupling chemistry with 4-(arylethynyl)phenylboronic acid derivatives. The diarylacetylene boron reagents were also synthesised by Pd-catalysed cross-coupling of a 4-bromophenyl boronic acid derivative. Two different organoboronic acid protecting strategies were tested – the boronic acid functionality could be converted to a trifluoroborate salt,[189] or protected as a boronate ester by condensation of the boronic acid with neopentyl glycol. The boronate ester

133

protecting group strategy was found to be more useful as the products were more easily purified (organotrifluoroborate salts are unstable to silica gel).[110] The boronate ester products were coupled to 7-iodo-7-deazaadenine nucleosides by aqueous Pd-catalysed Suzuki-Miyaura coupling, employing the water-soluble ligand TPPTS (Scheme 6.1). The nucleoside products were isolated in good yields.

X = OH	Ar = Ph-, 74%
	Ar = 4-MeO-C$_6$H$_4$-, quant.
	Ar = 4-MeO-C$_6$H$_4$-, 71%
	Ar = 3-Thienyl-, 77%
	Ar = 4-CF$_3$-C$_6$H$_4$-, 68%
X = H	Ar = Ph, quant.
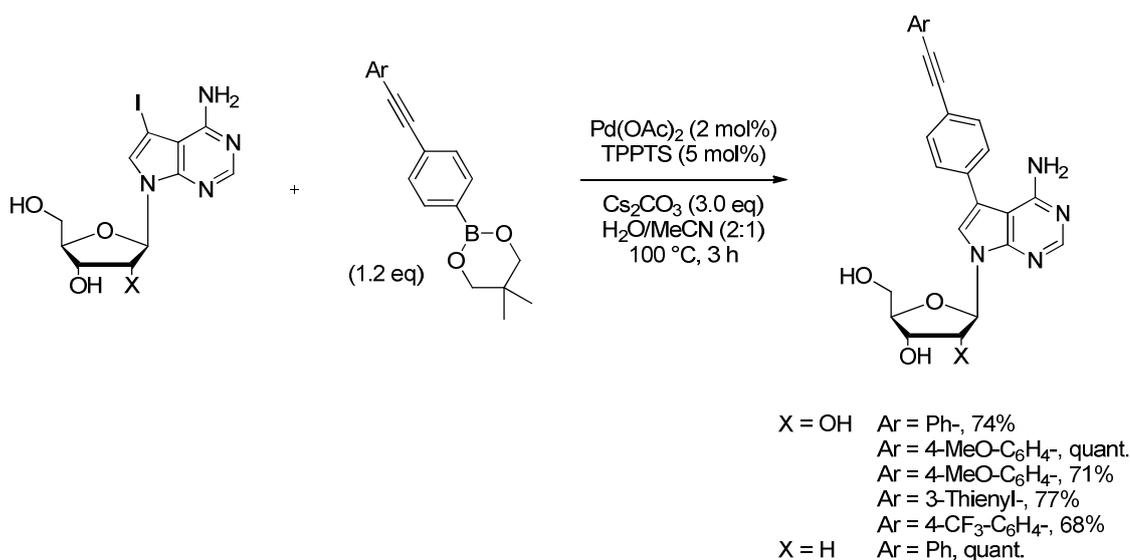
Scheme 6.1 Cross-coupling of 7-modified-7-deazaadenosines with organoboronate esters.

6.2 Characterisation of the 7-deazaadenosine analogues by UV-Vis absorption and fluorescence spectroscopy

Photophysical characterisation of the *C*-modified nucleosides exhibited promising UV-vis and fluorescence properties. In DMSO, the 7-(4-(arylethynyl)phenyl)-7-deazaadenosines had high quantum yields (up to 0.78), and absorbance maxima shifted away from the intrinsic absorbances associated with proteins and nucleic acids. These quantum yields are a significant improvement on other reported fluorescent 7-deazadenine nucleosides. The nucleosides with smaller π-systems, such as phenyl and phenylethynyl substituents, were found to be only very weakly fluorescent, with low quantum yields (<0.02). The 4-CF$_3$-substituted analogue (**20e**) has a larger Stokes' shift than the other analogues, which could be due to the 'push-pull' effect of the electron-withdrawing CF$_3$ group and the electron-rich heterocycle. Investigation of nucleosides with a wider range of substitutions would be

required to confirm this. Solvatochromic studies showed a correlation between solvent polarity and Stokes' shift for the parent compound. Significant fluorescence quenching was observed for the parent compound in water, as was expected for these compounds, with the quantum yield decreasing from 0.74 to 0.15. Although the quantum yield in water is significantly lower, this is still a significant improvement on the fluorescent 7-deazaadenosines developed by Seela and others.[27] During the preparation of this work, Hocek and co-workers reported a similar system employing biaryl substituents with similar photophysical properties to the diarylacetylene compounds (including the observed decrease in quantum yield in water).[58] There is also a significant change in the fluorescence lifetime. In both DMSO and water the compound shows a double exponential decay, however in DMSO the major component (92%) is longer at 2.2 ns. In water, the ratio of the two components changes, and the lifetimes are decreased (0.4 ns (62%), 0.9 ns (38%)). The solvent-dependency of the fluorescence lifetimes could be further investigated in order to determine any correlation between polarity and the observed lifetimes. The changes in lifetime may provide a useful probe of chemical environment.

6.3 Characterisation of the modified oligonucleotides by fluorescence and CD spectroscopy

The parent nucleoside **20a** was incorporated into a series of short deoxyoligonucleotides (13mers) using solid-phase synthesis. The photophysical properties were found to be significantly different in the oligonucleotides, with the absorbance maximum at 322 nm and the emission maximum at *ca.* 400 nm, irrespective of sequence. The quantum yields, however, vary from 0.024 to 0.237, depending on the sequence (Table 6.1) but with no direct correlation to the neighbouring nucleotides. Many fluorescent nucleotides are quenched upon incorporation into oligonucleotides,[164] so the conservation of the quantum yields in this case is highly beneficial. Acrylamide quenching experiments were used to probe the structure of the single-strand oligonucleotides further, and in combination with examination of the temperature-dependency of the fluorescence spectra, indicated that the nature of the secondary structure of the oligonucleotides varied between sequences. Some correlation was observed between solvent accessibility (as measured by Stern-Volmer plots) and quantum yield, with the most solvent-accessible fluorescent nucleotide (**ODN3**) having the highest quantum yield. Circular dichroism spectroscopy also suggested

variation in the secondary structure of the single-strand oligonucleotides, with **ODN1** (the least solvent accessible fluorophore) showing very little ordered structure, meaning that it may be randomly coiled in such that the fluorophore is buried. This is supported by the changes in fluorescence intensity observed upon heating – **ODN1** shows an increase in fluorescence intensity as the temperature increased. The other oligonucleotides, however, show significant helical structure consistent with the formation of an ordered structure similar to B-form DNA, except **ODN3,** which is more consistent with an A-type structure. Further comparison of the CD spectra of the unmodified single-strand oligonucleotides with those recorded for the modified oligonucleotides would help elucidate whether the differences in secondary structure are due to the presence of the modified nucleotide.

Table 6.1 Summary of the fluorescence properties of the single- and double-stranded oligonucleotides.

	Sequence	λ_{em} ssDNA^{a} /nm	λ_{em} dsDNA^{a} /nm	Φ ssDNA	Φ dsDNA
ODN1	$^{5'}$d(CGCTTG**a**GTTCGC)$^{3'}$	400	412	0.025	0.110
ODN2	$^{5'}$d(CGCTTC**a**CTTCGC)$^{3'}$	400	409	0.024	0.035
ODN3	$^{5'}$d(CGCAAG**a**GAACGC)$^{3'}$	405	412	0.237	0.154
ODN4	$^{5'}$d(CGCAAC**a**CAACGC)$^{3'}$	404	410	0.086	0.054
ODN5	$^{5'}$d(CGCAAT**a**TAACGC)$^{3'}$	405	400	0.061	0.026

^{a} λ_{ex} = 322 nm.

The single-strand oligonucleotides were hybridised with their complementary strands, and the effect of the modified nucleotide was measured by comparison of the melting temperatures with those of the unmodified duplexes. Significant destabilisation was observed for the modified duplexes, although not enough to suggest the absence of a base-pair. Unlike the single-strand oligonucleotides, the fluorescence properties of the duplexes have a clearer correlation with the sequence. The emission maximum and quantum yield (as well as the fluorescence intensities) are dependent on the bases neighbouring the fluorescent nucleotide. Those sequences with neighbouring guanosines have the highest fluorescence. When the *C*-modified nucleotide is between pyrimidine bases, some

fluorescence quenching is observed, with greater quenching observed for the sequence with thymidines (**ODN5**). Examination of differences between the single-strand and duplexed oligonucleotides was not found to be informative, most likely because the single-strand oligonucleotides do not have comparable secondary structures (see above). CD spectroscopy of the duplexes showed them to be consistent with B-form DNA, and very similar to the unmodified duplexes. Single-crystal X-ray diffraction of the duplexes would be invaluable for identification of any conformational or structural changes caused by introduction of the modified nucleotide.

6.4 Characterisation of 7-deazaadenosine analogue as a base-discriminating fluorescent probe

The application of the fluorescent nucleotide as a probe of hybridisation (base-pair formation) was also investigated. Complementary strands with different bases opposite the fluorescent nucleotide were used to examine the effect of base-pair mismatches. Comparison with the correctly matched duplex indicated that cytidine mismatches cause significant fluorescence quenching. No clear pattern was observed for purine mismatched pairs. Further work into the origins of the mismatch discrimination between thymidine and cytidine led to investigation of the effect of pH on the oligonucleotides. It was found that the mismatch discrimination is due to formation of a protonated $A^+ \cdot C$ mismatch base-pair as previously proposed by Seela and co-workers (Figure 6.2).[173] When a 7-deazaadenosine-cytidine mismatch is present, the pK_a of the 7-deazaadenosine is increased, so that the adenosine analogue is partially protonated at pH 7 (Figure 6.3). As observed for the free nucleotide, protonation of the nucleotide heterocycle caused fluorescence quenching. Examination of the fluorescence lifetimes across the range of pH values would add further understanding to the quenching effect.
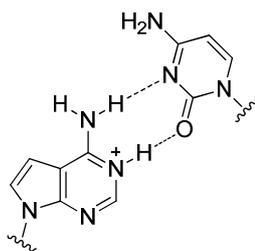
Figure 6.2 c^{7}A$^+\cdot$C wobble pair proposed by Seela and co-workers.
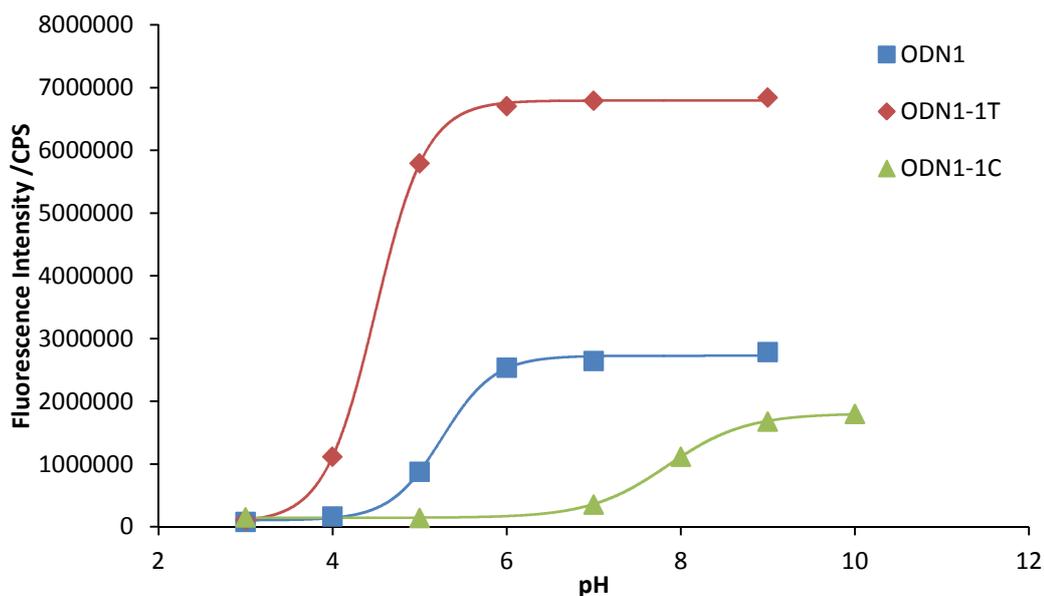
Figure 6.3 Effect of varying pH on the fluorescence emission at 400 nm of **ODN1** and its duplexes (2.5 µM).

Further work could examine the optimum pH for mismatch discrimination, which may be lower than pH 7. As discussed above, there is no clear pattern to the fluorescence intensities observed for the purine mismatches. There are two most likely reasons for the variation observed: (a) it is possible that the mismatches cause subtle conformational changes which change the solvent exposure of the residues (this could be further investigated by acrylamide quenching studies of all the duplexes with purine mismatches); (b) slight changes in the pK$_a$ of the nucleotides could partially protonate the heterocycle, causing some fluorescence quenching. Examination of the pH dependence of the fluorescence spectra, and calculation of the pK$_a$ of the nucleotide in each example would confirm or disprove this possibility.

6.5 Development of milder conditions for the direct arylation of 2′-deoxyadenosine

In this second strand to the project, efforts were made to improve the direct arylation of 2′-deoxyadenosine, so that it could be used under milder conditions that in the future would be more compatible with more complex oligonucleotide systems. After screening a variety of Pd^0 catalysts in the reaction at 60 °C, optimum yields were obtained when Pd(PVP) nanoparticles (*ca.* 2 nm) were used as the catalyst. The yield of 52% is lower than the reaction yield at 80 °C, but still respectable. Pd clusters or nanoparticles appear to play a key role in this reaction, but it is not clear whether they are acting as a reservoir of molecular Pd, or if the reaction takes place on the surface of the particle. More detailed studies, similar to those conducted for Suzuki-Miyaura couplings,[184] could help identify the catalytic species present (work that will continue in York over the coming years). This temperature is extremely low for this class of C-H functionalisation reaction, as most chemistry of this type is carried out at extremely high temperatures (120-160 °C). This lower temperature is unusual and emerging as a compatible system for use with more complex biomolecules. Unfortunately this is beyond the scope of this PhD project, but this would be the next step.

Further information about the role of excess CuI in these reactions has also been elucidated. It appears that CuI reacts with Cs_2CO_3 to release CO_2. This presumably proceeds *via* a Cu^I carbonate species. It is not clear with the current information if this is a catalytically active intermediate, or a decomposition pathway. Kinetic studies would help explain the mechanism of this reaction, however due to the low solubility of CuI and Cs_2CO_3 in the reaction solvent, as well as difficulties in accurate determination of the amount of starting material or product (the nucleosides bind to Cu^I), this is non-trivial.

6.6 Future research perspectives

A novel class of fluorescent base-discriminating 7-deazaadenine nucleosides has been synthesised using Pd-catalysed cross-coupling chemistry. The fluorescent nucleosides were incorporated into a series of oligonucleotides, and found to show promising base discrimination between pyrimidine match (T) and mismatch (C) base-pairs. Investigation of

the mechanism of the mismatch discrimination showed that protonation of the fluorescent nucleotide and formation of a wobble pair was the cause of the fluorescence quenching observed. The highly-pH dependent wobble pair means that mismatch discrimination is very pH sensitive. This has implications for other base-discriminating nucleotides reported in the literature, and more generally for the use of 7-deazaadenosines as adenosine analogues. Formation of non-Watson-Crick base-pairs may cause errors in transcription or replication of nucleic acids containing these residues. Further investigation into the application of these analogues in sequences associated with SNPs should assess the usefulness of these modified nucleotides for biomedical diagnostic applications. Due to the highly pH-sensitive nature of the fluorescent nucleotides, they could also have applications as pH-sensitive probes. The range of analogues synthesised could be expanded to include a wider range of substituents, particularly those with more polar (and therefore solubilising) properties. This may also simplify phosphorylation of the nucleotides, and allow investigations into their compatibility with nucleic acid polymerase enzymes. Another area for expansion would be the inclusion of 7-deazaguanine nucleosides containing the same 4-(arylethynyl)phenyl fluorophores. Studies into the direct arylation of 2'-deoxyadenosine has shed light on some key mechanistic details, but further studies would be required to fully understand the reaction mechanism. Application of this methodology to mismatch discrimination will be further investigated within our laboratories.

7   Experimental

General details

Dry, N_2-saturated solvents were collected from a Grubbs-type solvent system[190] in flame and vacuum-dried glassware. Extra Dry DMF was purchased from Acros and deoxygenated by bubbling with N_2 for 30 min. Methanol was dried according to the procedure outlined by Williams *et al.*[191] Pd(OAc)$_2$ was purchased from Precious Metals Online. 6-Chloro-7-deazapurine was purchased from Chemos GmbH. All other chemicals were purchased from Sigma-Aldrich or Alfa Aesar. Proton (^{1}H), carbon (^{13}C), phosphorus (^{31}P), fluorine (^{19}F) and boron (^{11}B) NMR were recorded using a Jeol ECX400, ECS400 or Bruker AMX500 spectrometer. ^{1}H NMR were referenced to undeuterated solvent according to Fulmer *et al.*[192] ^{13}C spectra were referenced to deuterated solvent. ^{31}P, ^{11}B and ^{19}F NMR were proton decoupled and reported relative to to 85% H_3PO_4, $BF_3.OEt_2$ and $CFCl_3$ respectively. In all cases, the *ipso* carbon of boron-containing compounds was not observed due to quadrupolar coupling. Mass spectrometry was recorded using ESI on a Bruker Daltronics micrOTOF machine, or EI on a Waters GCT Premier. Mass spectrometry of oligonucleotides was carried out using MALDI on a solariX FTMS 9.4T by dissolving the samples in water (*ca.* 500 μM) and mixing 1:1 with a solution of 3-hydroxypicolinc acid (0.5 M, 1:1 MeCN/H_2O). Melting points (Mp) were recorded using a Stuart digital SMP3 machine. UV/Vis spectra were recorded using a Jasco V-550 or a Jasco V-560 spectrophotometer. Oligonucleotide melting curves were measured with a Jasco V-550 fitted with a Jasco PTC-423 Peltier device. Fluorescence spectra were recorded using a Spectromax 3 fluorimeter. Quantum yields were measured using an integrating sphere. Circular dichroism spectra were run on a Jasco J810 CD Spectrophotometer fitted with a Jasco PTC-423 Peltier device. High Performance Liquid Chromatography (HPLC) was carried out using an Agilent 1100 system with DAD (diode array detector). Analytical HPLC was carried out on a Zorbax Eclipse XDB-C18 5 μm reverse-phase column, monitoring absorbance at 254 nm and semi-preparative HPLC was carried out on a Luna 5 μm C18(2) 100 Å column.

4-Chloro-5-iodo-7H-pyrrolo[2,3-d]pyrimidine 13
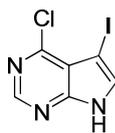
Based on the procedure reported by Wilulski *et al.*[104]

4-Chloro-7H-pyrrolo[2,3-d]pyrimidine (2.61 g, 17.1 mmol, 1.0 eq) and KOH (2.39 g, 42.7 mmol, 2.5 eq) were dissolved in DMF (30 ml). Iodine (4.39 g, 17.3 mmol, 1.01 eq) in DMF (30 ml) was added slowly, and the reaction was stirred at room temperature for 30 min. The reaction mixture was then poured onto water and ice (*ca.* 800 ml) containing 0.5 % ammonia and 0.1 % sodium metabisulfite. A white precipitate formed slowly, and this was filtered off. The precipitate was dissolved in EtOAc, dried over sodium sulfate and filtered. The volatiles were removed *in vacuo* to yield the product as a white solid (3.53 g, 12.7 mmol, 74 %). Spectroscopic characterisation was consistent with that previously reported.[103] ^{1}H NMR (400 MHz, DMSO-d_6) δ 12.95 (br s, 1H, NH), 8.60 (s, 1H, C2-H), 7.93 (d, *J* = 2.5, 1H, C6-H); ^{13}C NMR (101 MHz, DMSO-d_6) δ 151.54, 150.77, 150.52, 133.91, 115.81, 51.75; ESI HRMS [MH]$^+$ 279.9134 (Calcd. for $C_6H_4ClIN_3$ 279.9133).

4-Chloro-5-iodo-7-[(2,3,5-tri-*O*-benzoyl)-β-D-ribofuranosyl]-7H-pyrrolo[2,3-d]pyrimidine 14
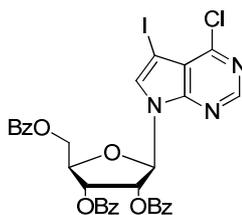
Based on the procedure reported by Seela and Ming.[103]

4-Chloro-5-iodo-7H-pyrrolo[2,3-d]pyrimidine (729 mg, 2.61 mmol, 1.0 eq) was stirred with BSA (0.77 ml, 3.1 mmol, 0.5 eq) in dry acetonirile (20 ml) at room temperature for 15 min. 1-*O*-Acetyl-2,3,5-*O*-benzoyl-β-D-ribofuranose (1.45 g, 2.87 mmol, 1.1 eq) and TMSOTf (0.20 ml, 3.1 mmol, 0.5 eq) were added and the reaction mixture was stirred for a further 10 min at rt, followed by 1.5 h at 80 °C. The reaction mixture was allowed to cool to rt, then

diluted with EtOAc (40 ml). The reaction mixture was then washed with sat. aq. NaHCO$_3$ (2 x 30 ml) and brine (2 x 30 ml), dried over Na$_2$SO$_4$ and filtered. The volatiles were removed *in vacuo* to yield the crude product as a yellow gum. This was redissolved in cyclohexane/EtOAc (20 ml 1:1 v/v) and absorbed on to silica. It was purified by column chromatography on silica gel, eluting with cylcohexane/EtOAc (10:1 v/v) to yield the title compound as a white solid (1.20 g, 1.66 mmol, 71 %). Spectroscopic characterisation was consistent with that previously reported.[103] ^{1}H NMR (400 MHz, DMSO-d_6) δ 8.60 (s, 1H, C2-H), 8.30 (s, 1H, C6-H), 7.99 (dd, *J* = 8.4, 1.3 Hz, 2H, ArH), 7.94 (dd, *J* = 8.4, 1.3 Hz, 2H, Ar-H), 7.85 (dd, *J* = 8.4, 1.3 Hz, 2H, Ar-H), 7.71 – 7.61 (m, 3H, Ar-H), 7.57 – 7.39 (m, 6H, Ar-H), 6.71 (d, *J* = 5.1 Hz, 1H, C1'-H), 6.29 (dd, *J* = 6.1, 5.1 Hz, 1H, C2'-H), 6.13 (app t, *J* = 5.8 Hz, 1H, C3'-H), 4.86 (ddd, *J* = 5.8, 5.0, 3.9 Hz, 1H, C4'-H), 4.80 (dd, *J* = 12.1, 3.9 Hz, 1H, C5'-Ha), 4.68 (dd, *J* = 12.1, 5.0 Hz, 1H, C5'-Hb); ^{13}C NMR (101 MHz, DMSO-d_6) δ 166.0, 165.2, 165.0, 152.2, 151.5, 151.2, 134.8, 134.5, 134.5, 134.1, 130.0, 129.9, 129.9, 129.9, 129.8, 129.7, 129.7, 129.4, 129.3, 129.2, 129.2, 129.1, 128.8, 87.0, 79.8, 74.0, 71.2, 63.9, 55.1 (29 of 32 resonances observed); ESI HRMS [MH]$^+$ 724.0336 (Calcd. for C$_{32}$H$_{24}$ClIN$_3$O$_7$ 724.0342).

4-Amino-5-iodo-7-(β-ᴅ-ribofuranosyl)-7H-pyrrolo[2,3-d]pyrimidine 15
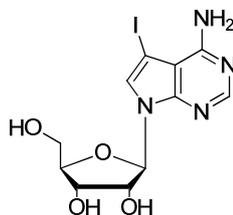
Based on the procedure reported by Seela and Ming.[103]

4-Chloro-5-iodo-7-[(2',3',5'-tri-*O*-benzoyl)-β-ᴅ-ribofuranosyl]-7H-pyrrolo[2,3-d]pyrimidine (362 mg, 0.500 mmol) was added to a sealed tube and dissolved in aq. NH$_3$ (5 ml, 25 %) and dioxane (5 ml). The reaction was stirred at 60 °C for 3 days. The volatiles were removed *in vacuo* and the crude product was redissolved in methanol and absorbed on to silica. The product was then purified by column chromatography on silica gel eluting with CH$_2$Cl$_2$/MeOH/Et$_3$N (95:4:1). The product was dried *in vacuo* to yield 4-amino-5-iodo-7-(β-ᴅ-ribofuranosyl)-7H-pyrrolo[2,3-d]pyrimidine as a white solid (171 mg, 0.436 mmol, 87 %). Spectroscopic characterisation was consistent with that previously reported.[103] ^{1}H NMR

(400 MHz, DMSO-d_6) δ 8.10 (s, 1H, C2-H), 7.67 (s, 1H, C6-H), 6.67 (br s, 2H, NH$_2$), 6.02 (d, J = 6.4 Hz, 1H, C1'-H), 5.31 (d, J = 6.4 Hz, 1H, C2'-OH), 5.16 (app t, J = 6.0 Hz, 1H, C5'-OH), 5.12 (d, J = 5.0 Hz, 1H, C3'-OH), 4.35 (app dt, J = 6.4, 5.0 Hz, 1H, C2'-H), 4.06 (app dt, J = 5.0, 3.1 Hz, 1H, C3'-H), 3.88 (app q, J = 3.8 Hz, 1H, C4'-H), 3.61 (ddd, J = 12.0, 5.0, 4.0 Hz, 1H, C5'-Ha), 3.52 (ddd, J = 12.0, 6.0, 3.9 Hz, 1H, C5'-Hb); ^{13}C NMR (101 MHz, DMSO-d_6) δ 157.2, 151.9, 150.2, 127.2, 103.3, 86.8, 85.2, 73.9, 70.5, 61.6, 51.9; ESI HRMS [MH]$^+$ 393.0058 (Calcd. for C$_{11}$H$_{14}$IN$_4$O$_4$ 393.0054).

4-Amino-5-phenylethynyl-7-(β-D-ribofuranosyl)-7H-pyrrolo[2,3-d]pyrimidine 16
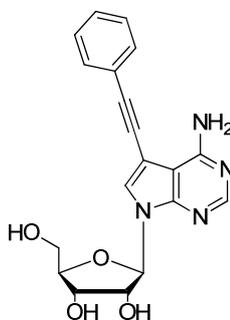
4-Amino-5-iodo-7-(β-D-ribofuranosyl)-7H-pyrrolo[2,3-d]pyrimidine (55 mg, 0.14 mmol), PdCl$_2$(PPh$_3$)$_2$ (0.9 mg, 1 μmol, 1 mol%), CuI (0.5 mg,3 μmol, 2 mol%) and were added to a Schlenk tube under a N$_2$ atmosphere. Triethylamine (59 μl, 0.42 mmol, 3 eq) and DMF (3 ml) were added and the reaction was then stirred at room temperature for 5 min. Phenylacetylene (19 μl, 0.17 mmol, 1.2 eq) was then added and the reaction was stirred for 4 hours at room temperature. The solvent was removed *in vacuo* and the resulting crude product was absorbed on to silica and purified by column chromatography (eluting with CH$_2$Cl$_2$/MeOH, 98:2 → 9:1 v/v). The product was dried *in vacuo* to give the product as a pale yellow solid (42 mg, 0.115 mmol, 83 %). Mp 149-150 °C (decomp.); ^{1}H NMR (400 MHz, DMSO-d_6) δ 8.15 (s, 1H, C2-H), 7.90 (s, 1H, C6-H), 7.64 − 7.55 (m, 2H, Ar-H), 7.48 − 7.38 (m, 3H, Ar-H), 6.73 (br s, 2H, NH$_2$), 6.06 (d, J = 6.1 Hz, 1H, C1'-H), 5.27 (br s, 3H, OH), 4.40 (app t, J = 5.6 Hz, 1H, C2'-H), 4.10 (dd, J = 4.9, 3.4 Hz, 1H, C3'-H), 3.91 (app q, J = 3.5 Hz, 1H, C4'-H), 3.65 (dd, J = 11.9, 3.7 Hz, 1H, C5'-Ha), 3.55 (dd, J = 11.9, 3.7 Hz, 1H, C5'-Hb); ^{13}C NMR (101 MHz, DMSO-d_6) δ 157.8, 152.8, 149.8, 131.1, 128.7, 128.5, 127.2, 122.5, 102.2, 94.7, 86.3,

83.0, 76.7, 74.1, 72.1, 70.5, 61.6; ESI HRMS [MH]$^+$ 367.1393 (Calcd. for C$_{19}$H$_{19}$N$_4$O$_4$ 367.1401).

4-Amino-5-[(1-naphthyl)ethynyl]-7-(β-D-ribofuranosyl)-7H-pyrrolo[2,3-d]pyrimidine 17
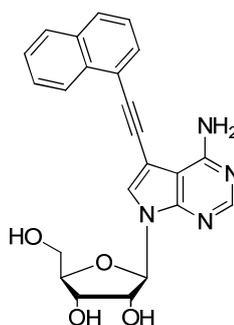
4-Amino-5-iodo-7-(β-D-ribofuranosyl)-7H-pyrrolo[2,3-d]pyrimidine (42 mg, 0.11 mmol), PdCl$_2$(PPh$_3$)$_2$ (0.7 mg, 1 μmol, 1 mol%) and CuI (0.4 mg, 2 μmol, 2 mol%) were added to a Schlenk tube under a N$_2$ atmosphere. Triethylamine (46 μl, 0.32 mmol, 3 eq), and DMF (2.5 ml) were added and the reaction was stirred at room temperature for 5 min. 1-Naphthylacetylene (19 μl, 0.13 mmol, 1.2 eq) was then added and the reaction was stirred overnight at room temperature. The solvent was removed *in vacuo* and the resulting crude product was absorbed on to silica and purified by column chromatography (eluting with CH$_2$Cl$_2$/MeOH, 98:2 → 9:1 v/v). The product was dried under reduced pressure to give the product as a pale yellow solid (40 mg, 0.095 mmol, 89 %). Mp 123-125 °C (decomp.); ^{1}H NMR (400 MHz, DMSO-d_6) δ 8.33 – 8.28 (d, J = 8.4, 1H, Ar-H), 8.16 (s, 1H, C2-H), 8.03 (s, 1H, C6-H), 8.02-7.95 (m, 2H, Ar-H), 7.88-7.84 (m, 1H, Ar-H), 7.68-7.63 (m, 1H, Ar-H), 7.62-7.57 (m, 1H, Ar-H), 7.56-7.51 (m, 1H, Ar-H), 6.79 (br s, 2H, NH$_2$), 6.06 (d, J = 6.1 Hz, 1H, C1′-H), 5.38 (d, J = 6.4 Hz, 1H, C2′-OH), 5.21 (app t, J = 5.5 Hz, 1H, C5′-OH), 5.15 (d, J = 4.8 Hz, 1H, C3′-OH), 4.41 (app q, J = 5.8 Hz, 1H, C2′-H), 4.11-4.07 (m, 1H, C3′-H), 3.90 (app q, J = 5.3 Hz, 1H, C4′-H), 3.64 (app dt, J = 11.7, 4.5 Hz, 1H, C5′-Ha), 3.54 (ddd, J = 11.7, 6.0, 3.9 Hz, 1H, C5′-Hb); ^{13}C NMR (101 MHz, DMSO-d_6) δ 157.6, 152.8, 149.9, 132.8, 132.3, 130.2, 128.8, 128.5, 127.5, 127.2, 126.7, 125.6, 125.3, 119.9, 102.1, 94.7, 89.0, 87.8, 87.1, 85.3, 73.9, 70.4, 61.5; ESI HRMS [MH]$^+$ 417.1543 (Calcd. for C$_{23}$H$_{21}$N$_4$O$_4$ 417.1557).

4-Amino-5-phenyl-7-(β-D-ribofuranosyl)-7H-pyrrolo[2,3-d]pyrimidine 18
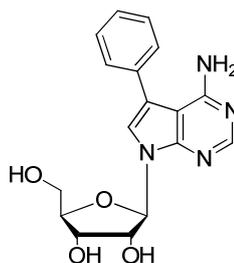
4-Amino-5-iodo-7-(β-D-ribofuranosyl)-7H-pyrrolo[2,3-d]pyrimidine (39.2 mg, 0.100 mmol, 1 eq), palladium acetate (1.2 mg, 5.0 µmol, 5 mol%), TPPTS (14.2 mg, 25.0 µmol, 25 mol%), phenylboronic acid (15.8 mg, 0.130 mmol, 1.3 eq) and sodium carbonate (32 mg, 0.30 mmol, 3 eq) were added to a Schlenk tube under a N_2 atmosphere. Water/acetonitrile (2:1, 1.2 ml) was added and the reaction mixture was stirred at 100 °C for 1.5 h. The reaction mixture was allowed to cool to room temperature, and the volatiles were removed *in vacuo*. The crude product was absorbed on to silica and purified by column chromatography, eluting with MeOH/CH_2Cl_2 (2:98 → 1:9 v/v). The product was obtained as a white solid (24.5 mg, 0.0720 mmol, 72 %). ^{1}H NMR (400 MHz, DMSO-d_6) δ 8.15 (s, 1H, C2-H), 7.55 (s, 1H, C6-H), 7.51 − 7.43 (m, 3H, Ar-H), 7.41 − 7.33 (m, 2H, Ar-H), 6.12 (d, *J* = 6.2 Hz, 1H, C1′-H), 5.35 (d, *J* = 6.4 Hz, 1H, C2′-OH), 5.20 (dd, *J* = 6.2, 5.0 Hz, 1H, C5′-OH), 5.14 (d, *J* = 4.8 Hz, 1H, C3′-OH), 4.49 − 4.43 (m, 1H, C2′-H), 4.18 − 4.03 (m, 1H, C3′-H), 3.90 (app q, J = 3.5 Hz, 1H, C4′-H), 3.63 (app dt, J = 11.9, 4.5 Hz, 1H, C5′-Ha), 3.53 (ddd, J = 11.9, 6.2, 3.7 Hz, 1H, C5′-Hb); ^{13}C NMR (101 MHz, DMSO-d_6) δ 157.3, 151.7, 150.9, 129.0, 128.5, 127.0, 121.2, 116.3, 100.5, 87.0, 85.1, 73.8, 70.6, 61.7; ESI HRMS [MH][+] 343.1403 (Calcd. for $C_{17}H_{19}N_4O_4$ 343.1401).

Potassium 4-bromophenyltrifluoroborate[109] 21
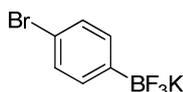
Prepared according to Molander *et al.*[109]

4-Bromophenylboronic acid (1.00 g, 5.00 mmol) was dissolved in THF (40 ml). Potassium hydrogen fluoride (2.30 g, 30.0 mmol, 6.0 eq) was dissolved in water (12.5 ml), and this was added slowly to the solution of boronic acid. The reaction mixture was stirred for 24 h at

room temperature. The solvents were then removed *in vacuo*, and the crude mixture was suspended in acetone (*ca.* 30ml). The inorganic salts were removed by filtration and washed with hot acetone (5 ml). The filtrate was collected and the solvent removed *in vacuo* to yield the product as a white solid (1.24 g, 4.73 mmol, 95 %). Spectroscopic characterisation was consistent with that previously reported. ^{1}H NMR (400 MHz, acetone-d_6) δ 7.40 (d, *J* = 7.8 Hz, 2H, Ar-H), 7.23 (d, *J* = 7.8 Hz, 2H, Ar-H); ^{13}C NMR (101 MHz, acetone-d_6) δ 134.7 (q, *J* = 1.7 Hz), 129.9, 119.7 (q, *J* = 0.9 Hz); ^{11}B NMR (128 MHz, acetone-d_6) δ 2.22 (br q, J_{F-B} = 50.4 Hz); ^{19}F NMR (376 MHz, acetone-d_6) δ -143.09 (q, J_{B-F} = 50.4 Hz); ESI MS *m/z* (%) 224 (99) [^{11}B^{81}BrM-K$^+$]$^-$, 223 (29) [^{10}B^{81}BrM-K$^+$]$^-$, 222 (100) [^{11}B^{79}BrM-K$^+$]$^-$, 221 (24) [^{10}B^{79}BrM-K$^+$]$^-$; HRMS [M-K$^+$]$^-$ 222.9549 (Calcd. for $C_6H_4BBrF_3$ 222.9548); Anal. calcd. for $C_6H_4BF_3BrK$: C 27.41, H 1.53; found: C 26.72, H 1.47.

Potassium 4-(phenylethynyl)phenyltrifluoroborate[108] 22a
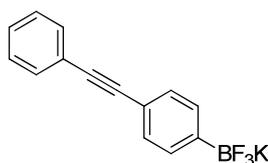
Prepared according to Kim *et al.*[108]

4-Bromophenyltrifluoroborate potassium (131 mg, 0.500 mmol), Pd(OAc)$_2$ (3.4 mg, 3 mol%) and XPhos (14.3 mg, 6 mol%) were added to a dry Schlenk tube under a N$_2$ atmosphere. DMSO (1 ml) and piperidine 74 µl, 0.75 mmol, 1.5 eq) were added, followed by phenylacetylene (82 µl, 0.75 mmol, 1.5 eq). The reaction was then stirred at 80 °C for 1 h. The volatiles were then removed under vacuum at 60 °C. The crude mixture was then resuspended in toluene (5 ml) by sonication and filtered. The precipitate formed was dried on the sinter, then washed into a clean flask with acetone. This second filtrate was collected and the solvent removed *in vacuo* to give a crude product. This was used without further purification. Spectroscopic characterisation was consistent with that previously reported. ^{1}H NMR (400 MHz, acetone-d_6) δ 7.56 – 7.47 (m, 4H, Ar-H), 7.41 – 7.34 (m, 3H, Ar-H), 7.30 (d, *J* = 7.5 Hz, 2H, Ar-H); ^{13}C NMR (101 MHz, acetone-d_6) 132.6 (br), 132.1 (br), 130.4 (br), 129.4 (br), 128.8 (br), 124.9, 120.2, 91.9, 88.1; ^{11}B NMR (128 MHz, acetone-d_6) δ

2.37 (br s); ^{19}F NMR (376 MHz, acetone-d_6) δ -139.54 (br s); ESI⁻ HRMS [M-K⁺]⁻ 245.0757 (Calcd. for $C_{14}H_9BF_3$ 245.0757).

Potassium 4-[(4-methoxyphenyl)ethynyl]phenyltrifluoroborate[108] 22b
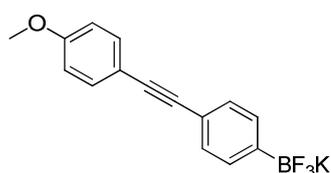
Prepared according to Kim *et al.*[108]

Potassium 4-bromophenyltrifluoroborate (39.3 mg, 0.150 mmol, 1.0 eq), Pd(OAc)$_2$ (1.0 mg, 3 mol%) and XPhos (4.3 mg, 6 mol%) were added to a dry Schlenk tube under a N$_2$ atmosphere. DMSO (0.3 ml) and piperidine (22 μl, 0.23 mmol, 1.5 eq) were added, followed by 4-ethynylanisole (29 μl, 0.23 mmol, 1.5 eq). The reaction was then stirred at 80 °C for 1 h. The volatiles were then removed under vacuum at 60 °C. The crude mixture was then resuspended in toluene (5 ml) by sonication and filtered. The precipitate formed was dried on the sinter, then washed into a clean flask with acetone. This second filtrate was collected and the solvent removed *in vacuo* to give a crude product. This was used without further purification. Spectroscopic characterisation was consistent with that previously reported. ^{1}H NMR (400 MHz, acetone-d_6) δ 7.50 (d, *J* = 7.7 Hz, 2H, Ar-H), 7.44 (d, *J* = 8.9 Hz, 2H, Ar-H), 7.26 (d, *J* = 7.7 Hz, 2H, Ar-H), 6.94 (d, *J* = 8.9 Hz, 2H, Ar-H), 3.83 (s, 3H, OCH$_3$); ^{13}C NMR (101 MHz, acetone-d_6) δ 160.5, 133.5, 132.6, 130.2, 120.6, 116.9, 115.0, 90.4, 88.1, 55.7 ^{11}B NMR (128 MHz, acetone-d_6) δ 2.41 (br s); ^{19}F NMR (376 MHz, acetone-d_6) δ -142.36 (br s); ESI⁻ MS *m/z* (%) 275 (100) [[11]BM-K⁺]⁻; 274 (23) [[10]BM-K⁺]⁻; HRMS [M-K⁺]⁻ 275.0865 (Calcd. for $C_{15}H_{11}BF_3O$ 275.0865).

(4-Methylsulfanylphenylethynyl)trimethylsilane
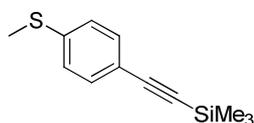
Prepared according to Sun *et al.*[193]

4-Bromothioanisole (508 mg, 2.50 mmol), [Pd(PPh$_3$)$_2$Cl$_2$] (17.5 mg, 1 mol%) and CuI (4.8 mg, 1 mol%) were dissolved in toluene in a flame-dried Schlenk tube. Triethylamine (25 ml) was added, followed by TMS-acetylene (1.24 ml, 3.00 mmol, 1.2 eq). The reaction mixture was stirred at 80 °C for 18 h. Water (20 ml) was added to the reaction mixture and the organic layer was separated. The aqueous layer was extracted with Et$_2$O (5 x 20 ml). The organic layers were combined, dried over MgSO$_4$, filtered and the solvent removed *in vacuo*. The crude oil was purified by column chromatography on silica gel eluting with Pet Ether 40-60 to give the product as a yellow oil (198 mg, 0.89 mmol, 36%). Spectroscopic characterisation was consistent with that previously reported. ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.37 (d, *J* = 8.7 Hz, 2H, Ar-H), 7.15 (d, *J* = 8.7 Hz, 2H, Ar-H), 2.48 (s, 3H, SCH$_3$), 0.24 (s, 9H, Si(CH$_3$)$_3$); ^{13}C NMR (101 MHz, CDCl$_3$) δ 139.7, 132.4, 125.8, 119.5, 105.0, 94.3, 15.5, 0.1; EI HRMS [M]$^+$ 220.0752 (Calcd. for C$_{12}$H$_{16}$SSi 220.0742).

4-Methylsulfanylphenylacetylene
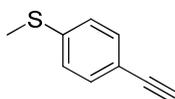
Prepared according to Tour *et al.*[194]

(4-Methylsulfanylphenylethynyl)trimethylsilane (290 mg, 1.33 mmol) was dissolved in MeOH/CH$_2$Cl$_2$ (40 ml, 1:1 *v/v*) under a N$_2$ atmosphere. K$_2$CO$_3$ (920 mg, 6.63 mmol) was added and the reaction was stirred at rt for 2 h. The reaction mixture was then poured onto water (40 ml) and extracted with Et$_2$O (4x40 ml). The volatiles were removed *in vacuo* to give the product as a yellow oil (146 mg, 0.989 mmol, 74%). Spectroscopic characterisation was consistent with that previously reported. ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.40 (d, *J* = 8.5 Hz, 2H, Ar-H), 7.17 (d, *J* = 8.9 Hz, 2H, Ar-H), 3.07 (s, 1H, CCH), 2.48 (s, 3H, SCH$_3$); ^{13}C NMR (101 MHz, CDCl$_3$) δ 140.2, 132.5, 125.9, 118.5, 83.6, 77.3, 15.5; APCI MS *m/z* (%) 133.0 (100) [MH]$^+$, 118.0 (1) [MH-Me]$^+$.

Potassium 4-[(4-methylsulfanylphenyl)ethynyl]phenyltrifluoroborate 22c
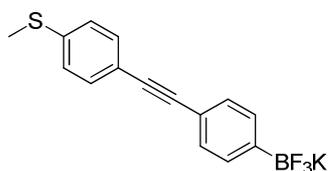
Potassium 4-bromophenyltrifluoroborate (39.3 mg, 0.150 mmol), Pd(OAc)$_2$ (1.0 mg, 3 mol%) and XPhos (4.3 mg, 6 mol%) were added to a dry Schlenk tube under a N$_2$ atmosphere. DMSO (0.3 ml) and piperidine (22 µl, 0.23 mmol, 1.5 eq) were added, followed by 4-ethynylthioanisole (33.4 mg, 0.225 mmol, 1.5 eq). The reaction was then stirred at 80 °C for 1 h. The volatiles were then removed under vacuum at 60 °C. The crude mixture was then resuspended in acetone (5 ml) and filtered through Celite. The filtrate was collected and the solvent removed *in vacuo*. The crude product was then resuspended in Et$_2$O and filtered. The precipitate was dried under vacuum to give the crude product. This was used without further purification. ^{1}H NMR (400 MHz, acetone-d_6) δ 7.51 (d, J = 8.1 Hz, 2H, Ar-H), 7.43 (d, J = 8.6 Hz, 2H, Ar-H), 7.28 (d, J = 8.1 Hz, 2H, Ar-H) 7.27 (d, J = 8.6 Hz, 2H, Ar-H) 2.52 (s, 3H, SCH$_3$); ^{13}C NMR (101 MHz, acetone-d_6) δ 132.7, 132.7, 132.5, 130.2, 126.7, 100.8, 92.7, 15.1 (8 of 11 resonances observed); ^{11}B NMR (128 MHz, acetone-d_6) δ 2.25 (br s); ^{19}F NMR (376 MHz, acetone-d_6) δ -147.17 (br s); ESI$^-$ MS m/z (%) 291 (100) [^{11}BMH]$^+$, 290 (24) [^{10}BMH]$^+$, HRMS [M-K$^+$]$^-$ 291.0639 (Calcd. for C$_{15}$H$_{11}$BF$_3$S 291.0635).

4-Bromophenylboronic acid neopentyl glycol ester {IUPAC name: 2-(4-Bromophenyl)-5,5-dimethyl-1,3,2-dioxaborinane} **26**
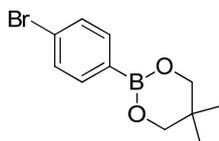
Prepared according to Zheng *et al.*[111]

4-Bromophenylboronic acid (1.00 g, 5.00 mmol, 1.0 eq) and 2,2,-dimethylpropanediol (573 mg, 5.50 mmol, 1.1 eq) were dissolved in toluene (50 ml) and heated to reflux in a Dean-

Stark apparatus for 16 h (or until production of water ceased). The toluene was removed *in vacuo*. The product was purified by column chromatography on silica gel, eluting with EtOAc/Pet Ether 40-60 (1:4 *v/v*). The product was isolated as a white solid (1.31 g, 4.87 mmol, 97%). Spectroscopic characterisation was consistent with that previously reported. Mp 107-110 °C (lit. 109-111 °C[195]); ^{1}H NMR (400 MHz, acetone-d_6) δ 7.67 (d, *J* = 8.4 Hz, 2H, Ar-H), 7.53 (d, *J* = 8.4 Hz, 2H, Ar-H), 3.80 (s, 4H, (CH$_2$)$_2$), 1.01 (s, 6H, (CH$_3$)$_2$); ^{13}C NMR (101 MHz, acetone-d_6) δ 136.5, 131.5, 125.8, 72.8, 32.4, 21.8; ^{11}B NMR (128 MHz, acetone-d_6) δ 25.6 (s); EI MS *m/z* (%) 267 (16) [^{10}B^{79}BrM]$^+$; 268 (100) [^{11}B^{79}BrM]$^+$; 269 (33) [^{10}B^{81}BrM]$^+$; 270 (97); 271 (12) [^{10}B^{81}BrM]$^+$; HRMS [M]$^+$ 268.0276 (Calcd. For C$_{11}$H$_{14}$O$_2$BBr 268.0272); Anal. calcd. for C$_{11}$H$_{14}$O$_2$BBr: C 49.13, H 5.25; found: C 48.88, H 5.21.

4-(Trimethylsilylethynyl)phenylboronic acid neopentyl glycol ester {IUPAC name: 2-(4-{Trimethylsilylethynyl}phenyl)-5,5-dimethyl-1,3,2-dioxaborinane} **28**
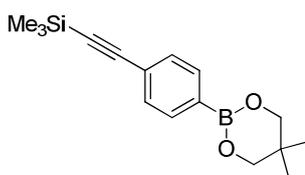
Prepared according to Zheng *et al.*[111]

4-Bromophenylboronic acid neopentyl glycol ester (1.08 g, 4.00 mmol, 1.0 eq), PdCl$_2$(PPh$_3$)$_2$ (140 mg, 5 mol%), CuI (38.0 mg, 5 mol%) and PPh$_3$ (210 mg, 20 mol%) were added to a microwave vial which was sealed with a septum and flushed with argon. DMF (2.2 ml) and Et$_2$NH (6.5 ml were added and the mixture was sparged with argon and stirred to produce a dark orange solution. Trimethylsilylacetylene (630 µl, 4.4 mmol, 1.1 eq) was added and the reaction mixture formed a pale yellow solution. After <2 min, a white precipitate (presumably Et$_2$NH$_2$I) was formed. The septum was replaced with a microwave vial cap and the mixture was heated to 120 °C in a microwave for 20 min. The dark orange/brown reaction mixture was then allowed to cool to room temperature and the volatiles removed *in vacuo*. The crude product was then suspended in water (25 ml) and extracted with CH$_2$Cl$_2$ (4x25 ml). The combined organic fractions were washed with brine (15 ml) and dried over Na$_2$SO$_4$, filtered and the solvent removed *in vacuo*. This crude material was purified by column chromatography on silica gel, eluting with Pet Ether 40-60, followed by

Et$_2$O/Pet Ether 40-60 (1:49 *v/v*) to yield the product as a white solid (804 mg, 2.80 mmol, 70%). Spectroscopic characterisation was consistent with that previously reported. Mp 94-98 °C; ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.72 (d, *J* = 8.1 Hz, 2H, Ar-H), 7.4 (d, *J* = 8.1 Hz, 2H, Ar-H), 3.76 (s, 4H, (CH$_2$)$_2$), 1.02 (s, 6H, (CH$_3$)$_2$), 0.25 (s, 9H, Si(CH$_3$)$_3$); ^{13}C NMR (101 MHz, CDCl$_3$) δ 133.7, 131.2, 125.2, 105.6, 100.1, 95.2, 72.5, 32.0, 22.1, 0.1; ^{11}B NMR (128 MHz, CDCl$_3$) δ 26.4 (br s); EI MS *m/z* (%) 270 (36) [^{10}BM-Me]$^+$; 271 (100) [^{11}BM-Me]$^+$; 285 (7) [^{10}BM]$^+$; 286 (25) [^{11}BM]$^+$; HRMS [M]$^+$ 286.1564 (Calcd. For C$_{16}$H$_{23}$O$_2$SiB 286.1558); Anal. calcd. for C$_{16}$H$_{23}$O$_2$SiB: C 66.13, H 8.10; found: C 66.17, H 8.31.

4-Ethynylphenylboronic acid neopentyl glycol ester {IUPAC name: 2-(4-Ethynylphenyl)-5,5-dimethyl-1,3,2-dioxaborinane} **29**
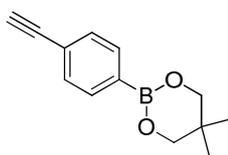
Prepared according to Zheng *et al.*[111]

4-(Trimethylsilylethynyl)phenylboronic acid neopentyl glycol ester (301 mg, 1.05 mmol, 1.0 eq) was dissolved in THF (10.5 ml) and cooled to 0 °C. TBAF (1.05 ml, 1.0 M in THF, 1.0 eq) was added and the reaction mixture stirred at 0 °C for 1 h. The mixture was then allowed to warm to r.t. and stirred for a further 2 h. Saturated NaHCO$_3$ solution (10 ml) was added and the mixture extracted with Et$_2$O (3x10 ml). The organic fractions were dried over Na$_2$SO$_4$, and the solvent removed *in vacuo*. The crude mixture could be used without further purification, or filtered through a short silica plug (ca. 2 cm) eluting with EtOAc/Pet Ether 40-60 (1:4 *v/v*) to give the title compound quantitatively as a white solid. Spectroscopic characterisation was consistent with that previously reported. Mp 75-77 °C (lit. 73-75 °C[196]) ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.75 (d, *J* = 8.2 Hz, 2H, Ar-H), 7.48 (d, *J* = 8.2 Hz, 2H, Ar-H), 3.77 (s, 4H, (CH$_2$)$_2$), 3.13 (s, 1H, CCH), 1.02 (s, 6H, (CH$_3$)$_3$); ^{13}C NMR (101 MHz, CDCl$_3$) 133.8, 131.3, 124.2, 72.5, 32.0, 22.0; ^{11}B NMR (128 MHz, CDCl$_3$) δ 25.8 (br s); EI MS *m/z* (%) 213.1143 (18) [^{10}BM]$^+$; 213.1794 (73) [^{11}BM-H]$^+$; 214 (100) [^{11}BM]$^+$; HRMS [M]$^+$ 214.1163 (Calcd. For C$_{13}$H$_{15}$O$_2$B 214.1168); Anal. calcd. for C$_{13}$H$_{15}$O$_2$B: C 72.94, H 7.06; found: C 72.94, H 7.17.

4-(Phenylethynyl)phenylboronic acid neopentyl glycol ester {IUPAC name: 2-(4-{Phenylethynyl}phenyl)-5,5-dimethyl-1,3,2-dioxaborinane} **27a**
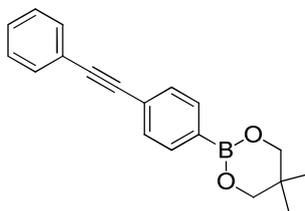
Prepared according to Zheng *et al.*[111]

4-Bromophenyl boronic acid neopentyl glycol ester (538 mg, 2.00 mmol, 1 eq), $PdCl_2(PPh_3)_2$ (70.2 mg, 5 mol%), CuI (19.0 mg, 5 mol%) and PPh_3 (105 mg, 20 mol%) were added to a microwave tube. A septum was used to seal the tube and it was flushed with argon. Diethylamine (3.2 ml) and DMF (1.1 ml) were added, followed by phenylacetylene (264 µl, 2.40 mmol, 1.2 eq). The septum was replaced by a lid and the reaction heated to 120 °C in a microwave (150 W) for 20 min. The solvent was then removed *in vacuo* to give the crude product, which was purified by column chromatography on silica gel eluting with Pet Ether 40-60, followed by EtOAc/Pet Ether 40-60 (1:4 *v/v*) to give the product as a pale yellow solid (508.3 mg, 88%). Spectroscopic characterisation was consistent with that previously reported. Mp 125-127 °C; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.78 (d, *J* = 7.8 Hz, 2H, Ar-H), 7.57 – 7.48 (m, 4H, Ar-H), 7.39-7.31 (m, 3H, Ar-H), 3.78 (s, 4H, $(CH_2)_2$), 1.03 (s, 6H, $(CH_3)_2$); ^{13}C NMR (101 MHz, $CDCl_3$) δ 133.8, 131.8, 130.8, 128.5, 128.4, 125.4, 123.4, 90.5, 89.9, 72.5, 32.1, 22.1; ^{11}B NMR (128 MHz, $CDCl_3$) δ 21.5 (br s); EI MS *m/z* (%) 178 (100) $[M-BO_2C_5H_{16}+H]^+$ 289 (5) $[^{10}BM]^+$; 290 (42) $[^{11}BM]^+$; HRMS $[M]^+$ 290.1443 (Calcd. For $C_{19}H_{19}BO_2$ 290.1482)); IR (ATR) $\tilde{v}$ (cm^{-1}) = 2901 (br w), 2503 (br), 2159 (br), 2029 (br), 1975, 1603, 1477, 1420, 1305, 1249, 1141, 1124.

4-([4-Methoxyphenyl]ethynyl)phenylboronic acid neopentyl glycol ester {IUPAC name: 2-(4-{[4-Methoxyphenyl]ethynyl}phenyl)-5,5-dimethyl-1,3,2-dioxaborinane} **27b**
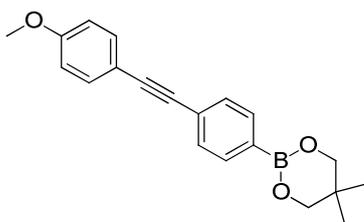
The title compound was prepared as described for **27a** from 4-bromophenyl boronic acid neopentyl glycol ester (40.3 mg, 0.150 mmol) and 4-ethynylanisole (23 μl, 0.18 mmol, 1.2 eq) and isolated as a yellow solid (29.8 mg, 0.0931 mmol, 62%). Mp 116-118 °C (decomp.); ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.76 (d, *J* = 8.3 Hz, 2H, Ar-H), 7.51 − 7.45 (m, 4H, Ar-H), 6.88 (d, *J* = 8.8 Hz, 1H, Ar-H), 3.83 (s, 3H, OCH$_3$), 3.77 (s, 4H, (CH$_2$)$_2$), 1.03 (s, 6H, (CH$_3$)$_2$); ^{13}C NMR (101 MHz, CDCl$_3$) δ 159.8, 133.8, 133.2, 130.7, 125.8, 115.6, 114.1, 90.5, 88.6, 72.5, 55.5, 32.1, 22.1; ^{11}B NMR (128 MHz, CDCl$_3$) δ 26.7 (br s); EI MS *m/z* (%) 319 (24) [[10]BM]$^+$; 320 (100) [[11]BM]$^+$; HRMS [M]$^+$ 320.1599 (Calcd. For C$_{20}$H$_{21}$BO$_3$ 320.1588); IR (ATR) $\tilde{v}$ (cm^{-1}) = 2959, 2918, 2211, 1599, 1476, 1419, 1305, 1245, 1141, 1089.

4-(3-Thienylethynyl)phenylboronic acid neopentyl glycol ester {IUPAC name: 2-(4-[3-Thienylethynyl]phenyl)-5,5-dimethyl-1,3,2-dioxaborinane} **27d**
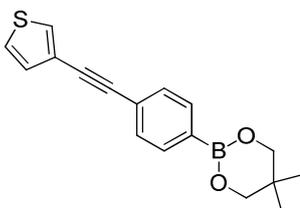
The title compound was prepared as described for **27a** from 4-bromophenyl boronic acid neopentyl glycol ester (107 mg, 0.500 mmol) and 3-ethynylthiophene (74 μl, 1.5 eq) and isolated as an off-white solid (101 mg, 0.341 mmol, 68%). Mp 139-142 °C (decomp.); ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.77 (d, *J* = 8.3 Hz, 2H, Ar-H), 7.52 (dd, *J* = 3.0, 1.2 Hz, 1H, Ar-H), 7.50 (d, *J* = 8.3 Hz, 2H, Ar-H), 7.30 (dd, *J* = 5.0, 3.0 Hz, 1H, Ar-H), 7.20 (dd, *J* = 5.0, 1.2 Hz, 1H, Ar-H), 3.77 (s, 4H, (CH$_2$)$_2$), 1.03 (s, 6H, (CH$_3$)$_2$); ^{13}C NMR (101 MHz, CDCl$_3$) δ 133.8, 130.7, 130.0, 128.8, 125.5, 125.4, 122.5, 89.4, 85.6, 72.5, 32.0, 22.1; ^{11}B NMR (128 MHz, CDCl$_3$) δ 26.4 (br s); EI MS *m/z* (%) 209 (13) [[10]BM-CHCHMe$_2$O]$^+$; 210 (50) [[11]BM-CHCHMe$_2$O]$^+$; 295

(21) $[^{10}BM]^+$; 296 (100) $[^{11}BM]^+$; HRMS $[M]^+$ 296.1050 (Calcd. For $C_{17}H_{17}SO_2B$ 296.1046); IR (ATR) $\tilde{v}$ (cm^{-1}) = 3095, 2949, 2897, 2160 (br), 2031 (br), 1976 (br), 1602, 1474, 1418, 1336, 1303, 1246, 1134, 1018.

4-([4-Trifluoromethylphenyl]ethynyl)phenylboronic acid neopentyl glycol ester {IUPAC name: 2-(4-{[4-Trifluoromethylphenyl]ethynyl}phenyl)-5,5-dimethyl-1,3,2-dioxaborinane} **27e**
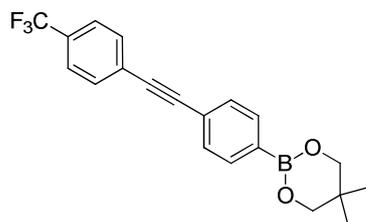
4-Ethynylphenyl boronic acid neopentyl glycol ester (107 mg, 0.500 mmol, 1 eq), PdCl$_2$(PPh$_3$)$_2$ (17.5 mg, 5 mol%), CuI (4.8 mg, 5 mol%) and PPh$_3$ (23.6 mg, 20 mol%) were added to a microwave tube. A septum was used to seal the tube and it was flushed with Ar. Diethylamine (1.5 ml) was added, followed by 4-iodobenzotrifluoride (110 µl, 0.750 mmol, 1.5 eq). The septum was replaced by a lid and the reaction heated to 120 °C in a microwave (150 W) for 25 min. The solvent was then removed *in vacuo* to give the crude product. The resulting solid was then purified by column chromatography on silica gel eluting with Et$_2$O/Pet Ether 40-60 (5:95 *v/v*) to give the product as a pale yellow solid (165 mg, 0.460 mmol, 92%). Mp 156-159 °C; ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.80 (d, *J* = 8.1 Hz, 2H, Ar-H), 7.63 (d, *J* = 9.0 Hz, 2H, Ar-H), 7.60 (d, *J* = 9.0 Hz, 2H, Ar-H), 7.53 (d, *J* = 8.1 Hz, 2H, Ar-H), 3.78 (s, 4H, (CH$_2$)$_2$), 1.03 (s, 6H, (CH$_3$)$_2$); ^{13}C NMR (101 MHz, CDCl$_3$) 133.9, 132.0, 131.0, 129.9 (q, *J* = 33 Hz), 127.3 (central peaks of q observed, *J* = 1.4 Hz), 125.4 (q, *J* = 4 Hz), 124.7 (central peaks of q observed, *J* = 272 Hz), 124.7, 122.7, 92.3, 89.0, 72.5, 32.1, 22.0; ^{11}B NMR (128 MHz, CDCl$_3$) δ 25.9 (br s); ^{19}F NMR (376 MHz, CDCl$_3$) δ -62.66 (s); EI MS *m/z* (%) 271 (21); 272 (88); 357 (21) $[^{10}BM]^+$; 358 (100) $[^{11}BM]^+$; HRMS $[M]^+$ 358.1301 (Calcd. For $C_{20}H_{18}BO_2F_3$ 358.1356); IR (ATR) $\tilde{v}$ (cm^{-1}) = 2954, 2916, 2214, 1739, 1608, 1477, 1422, 1316, 1249, 1104, 1064.

4-([4-Methylsulfanylphenyl]ethynyl)phenylboronic acid neopentyl glycol ester {IUPAC name: 2-(4-{[4-Methylsulfanylphenyl]ethynyl}phenyl)-5,5-dimethyl-1,3,2-dioxaborinane} **27c**
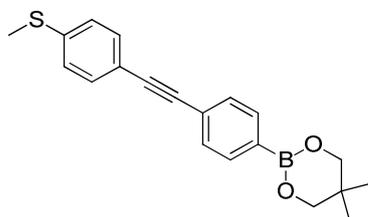
The title compound was prepared as described for **27e** from 4-ethynylphenyl boronic acid neopentyl glycol ester (32.1 mg, 0.150 mmol) and 4-bromothioanisole (45.7 mg, 1.5 eq) and isolated as a pale yellow solid (32.0 mg, 0.0952 mmol, 63%). Mp 148-150 °C; ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.77 (d, J = 8.1 Hz, 1H, Ar-H), 7.50 (d, J = 8.1 Hz, 2H, Ar-H), 7.44 (d, J = 8.4 Hz, 2H, Ar-H), 7.21 (d, J = 8.4 Hz, 2H, Ar-H), 3.78 (s, 4H, (CH$_2$)$_2$), 2.50 (s, 3H, SCH$_3$), 1.03 (s, 6H, (CH$_3$)$_2$); ^{13}C NMR (101 MHz, CDCl$_3$) δ 139.5, 133.8, 132.1, 130.8, 126.0, 125.5, 119.72, 119.71, 90.3, 90.0, 72.5, 32.1, 22.1, 15.5; ^{11}B NMR (128 MHz, CDCl$_3$) δ 26.7 (br s); EI MS m/z (%) 355 (21) [[10]BM]$^+$; 356 (100) [[11]BM]$^+$; HRMS [M]$^+$ 336.1359 (Calcd. For C$_{20}$H$_{21}$SO$_2$B 336.1359); IR (ATR) $\tilde{v}$ (cm^{-1}) = 2925 (br), 1730, 1603, 1477, 1421, 1342, 1309, 1250, 1128.

1-*O*-Methyl-2-deoxy-(α,β)-D-ribofuranose 31
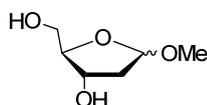
Prepared according to Adamo *et al.*[197]

2′-Deoxy-(α,β)-D-ribose (5.00 g, 37.2 mmol) was stirred in dry MeOH (50 ml) under a N$_2$ atmosphere. Acetyl chloride (265 μl, 10 mol%) was added and the reaction was stirred at ambient temperature for 1 h. The reaction was then quenched by addition of NaHCO$_3$ (2 g), and stirred for a further 5 min. The mixture was filtered through Celite, and the filtrate concentrated *in vacuo* to yield the product quantitatively as a yellow oil. The crude product was used without further purification. α:β 1:0.8 ^{1}H NMR (400 MHz, DMSO-d_6) δ 5.0 (br s),

4.97 (dd, J = 5.4, 2.5 Hz, H1β), 4.89 (dd, J = 5.7, 2.3 Hz, H1α), 4.65 (br s), 4.12 (br s), 4.09 (ddd, J = 10.1, 6.3, 4.6 Hz, H3β), 3.91 (app dt, J = 8.2, 5.0 Hz, H3α), 3.71-3.64 (m, H4α and H4β), 3.58-3.31 (m, H5α and H5β, partially obscured by H_2O), 3.23 (s, OMeα), 3.20 (s, OMeβ), 2.26 (ddd, J = 13.7, 8.2, 5.7 Hz, H2$_a$α), 1.98 (ddd, J = 13.4, 6.3, 2.6 Hz, H2$_a$β), 1.87 (app dt, J = 13.4, 5.4 Hz, H2$_b$β), 1.63 (ddd, J = 13.7, 5.0, 2.3 Hz, H2$_a$α); ^{13}C NMR (101 MHz, DMSO-d_6) δ 104.6, 104.1, 87.0, 85.0, 70.9, 70.1, 63.2, 61.4, 54.3, 54.3, 41.1, 41.0; ESI HRMS [MNa]$^+$ 171.0627 (Calcd. For $C_6H_{12}O_4Na$ 171.0628).

1-*O*-Methyl-2-deoxy-3,5,-di-O-*p*-toluoyl-(α,β)-D-ribofuranose 32
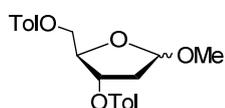
Prepared according to Rolland *et al.*[115]

1-*O*-Methyl-2-deoxy-(α,β)-D-ribofuranose was dissolved in dry pyridine (30 ml) under a N_2 atmosphere and cooled to 0 °C. *p*-Toluoyl chloride (11 ml, 2.6 eq) was added dropwise to the reaction at 0 °C with stirring. The reaction mixture was allowed to warm up to room temperature and stirred overnight. The pyridine was removed *in vacuo*. Water (75 ml) was added to the crude mixture, and the product was extracted with CH_2Cl_2 (3 x 75 ml). The organic layers were washed with sat. aq. $NaHCO_3$ (2 x 30 ml), HCl solution (2 M, 30 ml) and water (30 ml), then dried over $NaHCO_3$, filtered and the solvent removed *in vacuo*. The crude product could be used without further purification or purified by column chromatography on silica gel, eluting with EtOAc/Pet Ether 40-60 (1:9), to yield the product as a yellow gum (237 mg, 0.620 mmol, 83%). ^{1}H NMR (400 MHz, CDCl$_3$) δ 8.01-7.88 (m, Ar-H), 7.22 (t, J = 7.6 Hz, Ar-H), 5.62-5.57 (m, H1β), 5.41 (ddd, J = 8.3, 3.6, 2.2 Hz, H1α), 5.23 (dd, J = 5.5, 2.1 Hz, H4), 5.19 (d, J = 5.1 Hz, H4), 4.63-4.45 (m, H3α, H3β, H5α, H5β), 3.42 (s, OMe), 3.36 (s, OMe), 2.60-2.51 (m, C2α and C2β), 2.42 (s, ArMe), 2.41 (s, ArMe), 2.401 (s, ArMe), 2.397 (s, ArMe), 2.31 (app dt, J = 10.9, 5.8 Hz, H2), 2.19 (d, J = 13.8 Hz, H2); ^{13}C NMR (101 MHz, CDCl$_3$) δ 166.6, 166.4, 166.4, 166.2, 144.1, 144.0, 143.89, 143.8, 129.9, 129.9, 129.8, 129.8, 129.2, 129.2, 129.2, 129.2, 127.3, 127.2, 127.1, 127.0, 105.7, 105.1, 82.0, 81.0, 75.5, 74.7, 65.2, 64.4, 55.3, 55.2, 39.4, 39.4, 21.8, 21.8; ESI HRMS [MNa]$^+$ 407.1462 (Calcd. For $C_{22}H_{24}O_6Na$ 407.1465).

1-Chloro-2-deoxy-3,5,-di-O-*p*-toluoyl-α-D-ribofuranose 6
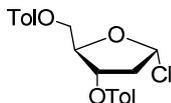
Prepared according to Rolland *et al.*[115]

Saturated HCl in acetic acid (*ca.* 2.3 M) was prepared according to the procedure by Kotera and co-workers.[198] 1-O-Methyl-2-deoxy-3,5,-di-O-*p*-toluoyl-(α,β)-D-ribofuranose (1.55 g, 4.03 mmol) was dissolved in acetic acid (11 ml). The resulting mixture was cooled to *ca.* 10 °C, and sat. HCl in AcOH (9 ml) was added. The mixture was stirred vigorously for a further 15 min, and a white precipitate was produced. The precipitate was collected by filtration and washed with cold Et$_2$O (50 ml) to yield the α-chloro sugar as a white solid (730 mg, 2.02 mmol, 50%).[**] Spectroscopic characterisation was consistent with that previously reported. ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.99 (d, *J* = 8.1 Hz, 2H, Ar-H), 7.89 (d, *J* = 8.1 Hz, 2H, Ar-H), 7.26 (d, *J* = 8.0 Hz, partially obscured by CHCl$_3$ peak, Ar-H), 7.24 (d, *J* = 8.4 Hz, 2H, Ar-H), 6.48 (app d, *J* = 5.3 Hz, 1H, C1′-H), 5.56 (ddd, *J* = 7.5, 3.0, 1.1 Hz, 1H, C3′-H), 4.86 (app dt, *J* = 4.2, 3.0 Hz, 1H, C′-H), 4.68 (dd, *J* = 12.1, 3.1 Hz, 1H, C5′-Ha), 4.59 (dd, *J* = 12.1, 4.2 Hz, 1H, C5′-Hb), 2.87 (ddd, *J* = 15.0, 7.4, 5.3 Hz, 1H, C2′-Ha), 2.75 (app dd, *J* = 15.1, 0.9 Hz, 1H, C2′-Hb), 2.42 (s, 3H, CH$_3$), 2.41 (s, 3H, CH$_3$), 2.0 (br s, 2H); ^{13}C NMR (101 MHz, CDCl$_3$) δ 166.5, 166.2, 144.4, 144.2, 130.0, 129.8, 129.3, 129.3, 126.9, 126.7, 95.4, 84.8, 73.6, 63.6, 44.6, 21.82, 21.78; ESI HRMS [M-Cl]$^+$ 353.1382 (Calcd. For C$_{21}$H$_{21}$O$_5$ 353.1384).

4-Chloro-5-iodo-7-[(3,5-di-*O*-*p*-toluoyl)-2-deoxy-β-D-ribofuranosyl]-7H-pyrrolo[2,3-d]pyrimidine[121] 33
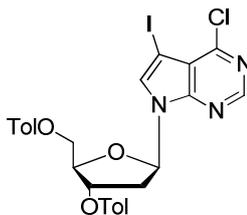
[**] Product isomerises in the presence of atmospheric water and in solution to give a mixture of α- and β-anomers. Stored in a vacuum desiccator over MgSO$_4$.

Prepared according to Seela *et al.*[121]

Finely powdered KOH (94.1 mg, 1.68 mmol, 2.1 eq) was stirred in dry MeCN (5 ml) under a N_2 atmosphere. TDA-1 (10 µl, 4 mol%) was added and the mixture was stirred for 5 min at rt. 6-Chloro-7-iodo-7-deazapurine (223 mg, 0.800 mmol, 1 eq) was added and the mixture was stirred for a further 5 min. 1-Chloro-2-deoxy-3,5,-di-O-*p*-toluoyl-α-D-ribofuranose (346 mg, 0.960 mmol, 1.2 eq) was added and the reaction was stirred for 20 min at rt. The crude mixture was absorbed on to silica gel and purified by column chromatography on silica gel, eluting with CH_2Cl_2/EtOAc (20:1) to yield the product as a yellow solid (370 mg, 0.580 mmol, 73%). Spectroscopic characterisation was consistent with that previously reported. ^{1}H NMR (400 MHz, DMSO-d_6) δ 8.64 (s, 1H, C2-H), 8.18 (s, 1H, C6-H), 7.94 (d, *J* = 8.2 Hz, 2H, Ar-H), 7.84 (d, *J* = 8.2 Hz, 2H, Ar-H), 7.35 (d, *J* = 8.2 Hz, 2H, Ar-H), 7.30 (d, *J* = 8.2 Hz, 2H, Ar-H), 6.75 (dd, *J* = 7.8, 6.5 Hz, 1H, C1′-H), 5.77 − 5.73 (m, 1H, C3′-H), 4.64 (app td, *J* = 8.1, 5.8 Hz, 1H, C4′-H), 4.58 − 4.50 (m, 2H, C5′-H), 3.09 (app dt, *J* = 14.3, 6.5 Hz, 1H, C2′-Ha), 2.76 (ddd, *J* = 14.3, 6.5, 2.6 Hz, 1H, C2′-Hb), 2.39 (s, 3H, CH_3), 2.37 (s, 3H, CH_3); ^{13}C NMR (101 MHz, DMSO-d_6) δ 165.5, 165.2, 151.3, 150.8, 150.6, 144.1, 143.8, 133.3, 129.5, 129.4, 129.3, 129.3, 126.5, 126.5, 116.8, 83.7, 81.6, 74.7, 64.0, 54.1, 36.2, 21.22, 21.20; ESI HRMS $[MH]^+$ 632.0438 (Calcd. For $C_{27}H_{24}ClIN_3O_5$ 632.0444).

4-Amino-5-iodo-7-(2-deoxy-β-D-ribofuranosyl)-7H-pyrrolo[2,3-d]pyrimidine 34
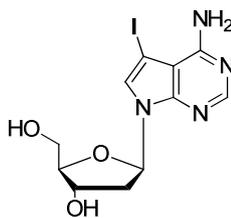
Prepared according to Seela *et al.*[199]

Compound **33** (2.00 g, 3.17 mmol) was suspended in an aqueous ammonia (25%)/dioxane mixture (1:1, 160 ml) and stirred in a sealed flask for 48 h at 110 °C. During this time the compound dissolved. The solvent was then removed *in vacuo* and the resulting mixture dissolved in MeOH/CH_2Cl_2 and absorbed onto silica gel. The product was purified by

column chromatography on silica gel eluting with CH_2Cl_2/MeOH (95:5 → 90:10) and isolated quantitatively as a white solid. Spectroscopic characterisation was consistent with that previously reported. ^{1}H NMR (400 MHz, DMSO-d_6) δ 8.10 (s, 1H, C2-H), 7.66 (s, 1H, C6-H), 6.67 (br s, 2H, NH$_2$), 6.48 (dd, J = 8.2, 5.9 Hz, 1H, C1′-H), 4.32 (app td, J = 6.1, 2.6 Hz, 1H, C3′-H), 3.80 (app td, J = 4.4, 2.6 Hz, 1H, C4′-H), 3.56 (app dt, J = 11.7, 5.0 Hz, 1H, C5′-Ha), 3.49 (ddd, J = 11.7, 5.7, 4.4 Hz, 1H, C5′-Hb), 2.45 (ddd, partially obscured by CD$_3$SOCD$_2$H, J = 13.1, 8.2, 5.7 Hz, 1H, C2′-Ha), 2.15 (ddd, J = 13.2, 5.9, 2.6 Hz, 1H, C2′-Hb); ^{13}C NMR (101 MHz, CDCl$_3$) δ 157.2, 151.9, 149.7, 126.8, 103.1, 87.4, 82.9, 70.9, 62.8, 61.9, 51.8; ESI HRMS [MH]$^+$ 377.0096 (Calcd. For $C_{11}H_{14}IN_4O_3$ 377.0105).

4-Amino-5-(4-bromophenyl)-7-(β-D-ribofuranosyl)-7H-pyrrolo[2,3-d]pyrimidine 19
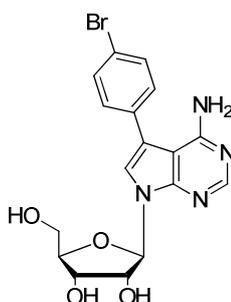
^{1}H NMR (400 MHz, DMSO-d_6) δ 8.11 (s, 1H, C2-H), 7.62 (d, J = 8.5 Hz, 2H, Ar-H), 7.55 (s, 1H, C6-H), 7.38 (d, J = 8.5 Hz, 2H, Ar-H), 6.20 (br s, 2H, NH$_2$), 6.07 (d, J = 6.4 Hz, 1H, C1′-H), 5.31 (d, J = 6.4 Hz, 1H, C2′-OH), 5.16 (dd, J = 6.1, 5.1 Hz, 1H, C5′-OH, 5.10 (d, J = 5.0 Hz, 1H, C3′-OH), 4.41 (app td, J = 6.4, 4.9 Hz, 1H, C2′-H), 4.06 (app q, J = 4.5 Hz, 1H, C3′-H), 3.86 (app q, J = 3.8 Hz, 1H, C4′-H), 3.59 (ddd, J = 11.9, 5.1, 3.8 Hz, 1H, C5′-Ha), 3.49 (ddd, J = 11.9, 6.1, 3.8 Hz, 1H, C5′-Hb); ^{13}C NMR (101 MHz, DMSO-d$_6$) δ 157.4, 151.8, 151.1, 133.7, 131.8, 130.4, 121.5, 120.0, 115.2, 100.2, 87.0, 85.1, 73.8, 70.6; MS (ESI) *m/z* (%) 423 (96) [[81]BrMH]$^+$, 421 (100) [[79]BrMH]$^+$, 291 (9) [[81]BrMH-ribose]$^+$, 289 (10) [[79]BrMH-ribose]$^+$; ESI HRMS [MH]$^+$ 421.0497 (Calcd. for $C_{17}H_{18}BrN_4O_4$ 421.0506).

General Procedure for synthesis of fluorescent analogues from organotrifluoroborates

4-Amino-5-iodo-7-(β-D-ribofuranosyl)-7H-pyrrolo[2,3-d]pyrimidine (49.0 mg, 0.125 mmol, 1.0 eq), Pd(OAc)$_2$ (0.6 mg, 2 mol%), TPPTS (2.8 mg, 4 mol%), sodium carbonate (66.3 mg,

0.625 mmol, 5.0 eq) and the potassium organotrifluoroborate (0.150 mmol, 1.2 eq)[†] were added to a Schlenk tube under a N$_2$ atmosphere. Water/acetonitrile (2:1, 3 ml) was added and the reaction mixture was stirred at 70 °C for 4 h. The reaction mixture was then neutralised with HCl (1 M), and the volatiles were removed *in vacuo*. The crude product was absorbed on to silica gel and purified by column chromatography on silica gel, eluting with AcOH/MeOH/CH$_2$Cl$_2$ (0.5:2.5:97 *v/v*).

General Procedure for synthesis of fluorescent analogues from organoboronate esters

4-Amino-5-iodo-7-(β-D-ribofuranosyl)-7H-pyrrolo[2,3-d]pyrimidine or 4-Amino-5-iodo-7-(2-deoxy-β-D-ribofuranosyl)-7H-pyrrolo[2,3-d]pyrimidine (1.0 eq), Pd(OAc)$_2$ (2 mol%), TPPTS (5 mol%), caesium carbonate (3.0 eq) and the organoboronate ester (1.2 eq) were added to a Schlenk tube under a N$_2$ atmosphere. Water/acetonitrile (2:1, *v/v*, 0.1 mM) was added and the reaction mixture was stirred at 100 °C for 3 h. The crude product was absorbed on to silica gel and purified by column chromatography, eluting with MeOH/CH$_2$Cl$_2$ (0:100 → 1:9 *v/v*) Small-scale reactions (0.050 mmol) were isolated by preparative TLC, eluting with 9:1 CH$_2$Cl$_2$/MeOH.

4-Amino-5-(4-[phenylethynyl]phenyl)-7-(β-D-ribofuranosyl)-7H-pyrrolo[2,3-d]pyrimidine 20a
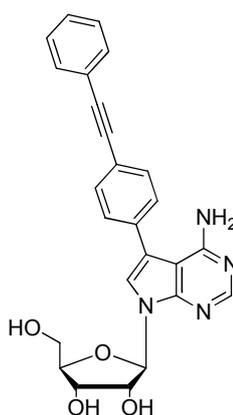
From organotrifluoroborate: The title compound was synthesised from 7-iodo-7-deazaadenosine (49.0 mg, 0.125 mmol) and potassium 4-

161

(phenylethynyl)phenyltrifluoroborate (0.15 mmol, 1.2 eq). The product was isolated as a white solid (31.0 mg, 0.070 mmol, 56%).

From organoboronate ester: The title compound was synthesised from 7-iodo-7-deazaadenosine (196 mg, 0.500 mmol) and 4-(phenylethynyl)phenylboronic acid neopentyl glycol ester (174 mg, 0.600 mmol, 1.2 eq). The product was isolated as a white solid (164 mg, 0.371 mmol, 74%).

^{1}H NMR (400 MHz, DMSO-d_6) δ 8.19 (br s, 1H, C2-H), 7.66 (d, J = 9.4 Hz, 3H, Ar-H), 7.61 – 7.50 (m, 4H, Ar-H), 7.46 – 7.42 (m, 3H, Ar-H), 6.28 (br s, 2H, NH$_2$), 6.13 (d, J = 6.2 Hz, 1H, C1'-H), 5.35 (d, J = 6.2 Hz, 1H, C2'-OH), 5.20 (t, J = 5.6 Hz, 1H, C5'-OH), 5.14 (d, J = 4.8 Hz, 1H, C3'-OH), 4.46 (app q, J = 6.2 Hz, 1H, C2'-H), 4.11 (m, 1H, C3'-H), 3.91 (app q, J = 3.9 Hz, 1H, C4'-H), 3.64 (ddd, J = 11.8, 5.6, 3.9 Hz, 1H, C5'-Ha), 3.53 (ddd, J = 11.8, 5.6, 3.9 Hz, 1H, C5'-Hb); ESI HRMS [MH]$^+$ 443.1706 (Calcd. for C$_{25}$H$_{23}$N$_4$O$_4$ 443.1714).

^{1}H NMR (400 MHz, CD$_3$OD) δ 8.16 (s, 1H, C2-H), 7.63 (d, J = 8.4 Hz, 2H, C), 7.56 – 7.52 (m, 3H), 7.50 (s, 1H), 7.41 – 7.35 (m, 3H), 6.11 (d, J = 6.4 Hz, 1H, C1'-H), 4.67 (dd, J = 6.4, 5.3 Hz, 1H, C2'-H), 4.31 (dd, J = 5.3, 2.8 Hz, 1H, C3'-H), 4.13 (q, J = 2.8 Hz, 1H, C4'-H), 3.87 (dd, J = 12.4, 2.8 Hz, 1H, C5'-Ha), 3.75 (dd, J = 12.4, 2.8 Hz, 1H, C5'-Hb); ^{13}C NMR (101 MHz, CD$_3$OD) δ 159.0, 152.4, 151.57, 135.8, 133.2, 132.5, 129.9, 129.59, 129.55, 124.5, 123.8, 123.5, 118.04, 102.9, 91.03, 90.99, 89.9, 87.3, 75.6, 72.5, 63.4.

4-Amino-5-(4-[{4-Methoxyphenyl}ethynyl]phenyl)-7-(β-D-ribofuranosyl)-7H-pyrrolo[2,3-d]pyrimidine 20b
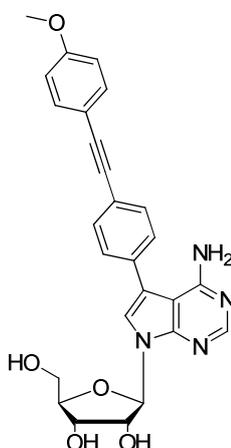
From organotrifluoroborate: The title compound was synthesised from 7-iodo-7-deazaadenosine (49.0 mg, 0.125 mmol) and potassium 4-[(4-methoxyphenyl)ethynyl]phenyltrifluoroborate (0.15 mmol, 1.2 eq). The product was obtained as a white solid (42.5 mg, 0.0900 mmol, 72 %).

From organoboronate ester: The title compound was synthesised from 7-iodo-7-deazaadenosine (19.6 mg, 0.0500 mmol) and 4-([4-methoxyphenyl]ethynyl)phenylboronic acid neopentyl glycol ester (19.2 mg, 0.600 mmol, 1.2 eq). The product was obtained as a white solid (23.5 mg, 0.050 mmol, quant.).

^{1}H NMR (400 MHz, CD$_3$OD) δ 8.16 (s, 1H, C2-H), 7.60 (d, J = 8.5 Hz, 2H, Ar-H), 7.52 (d, J = 8.5 Hz, 2H, Ar-H), 7.49 (s, 1H, C6-H), 7.46 (d, J = 9.0 Hz, 2H, Ar-H), 6.95 (d, J = 9.0 Hz, 2H, Ar-H), 6.11 (d, J = 6.3 Hz, 1H, C1'-H), 4.67 (dd, J = 6.3, 5.3 Hz, 1H, C2'-H), 4.30 (dd, J = 5.3, 2.9 Hz, 1H, C3'-H), 4.13 (app q, J = 2.9 Hz, 1H, C4'-H), 3.87 (dd, J = 12.3, 2.9 Hz, 1H, C5'-Ha), 3.83 (s, 3H, OCH$_3$), 3.75 (dd, J = 12.3, 2.9 Hz, 1H, C5'-Hb); ESI HRMS [MH]$^+$ 473.1806 (Calcd. for C$_{26}$H$_{25}$N$_4$O$_5$ 473.1819).

^{1}H NMR (500 MHz, DMSO-d_6) δ 8.18 (s, 1H, C2-H), 7.66-7.62 (m, 3H, Ar-H), 7-55-7.47 (m, 4H, Ar-H), 7.02 (d, J = 8.9 Hz, 2H, Ar-H), 6.24 (br s, 1H, NH$_2$), 6.15 (d, J = 6.2 Hz, 1H, C1'-H), 5.33 (d, J = 6.3 Hz, 1H, C2'-OH), 5.18 (app t, J = 5.6 Hz, 1H, C5'-OH), 5.12 (d, J = 4.9 Hz, 1H, C3'-OH), 4.48 (app dd, J = 11.6, 6.3 Hz, 1H, C2'-H), 4.13 (app dd, J = 8.2, 4.9 Hz, 1H, C3'-H), 3.93 (app q, J = 3.7 Hz, 1H, C4'-H), 3.66 (app dt, J = 12.1, 4.0 Hz, 1H, C5'-Ha), 3.60-3.53 (m, 1H, C5'-Hb); ^{13}C NMR (126 MHz, DMSO-d_6) δ 159.5, 157.3, 151.8, 151.1, 134.4, 132.9, 131.7, 128.5, 121.6, 120.8, 115.8, 114.4, 114.2, 90.0, 88.0, 87.0, 85.1, 73.8, 70.6, 61.6, 55.3.

4-Amino-5-(4-[{4-Methylsulfanylphenyl}ethynyl]phenyl)-7-(β-D-ribofuranosyl)-7H-pyrrolo[2,3-d]pyrimidine 20c
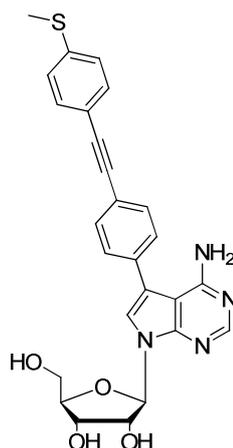
From organotrifluoroborate: The title compound was synthesised from 7-iodo-7-deazaadenosine (49.0 mg, 0.125 mmol) and potassium 4-[(4-methylsulfanylphenyl)ethynyl]phenyltrifluoroborate (0.15 mmol, 1.2 eq). The product was obtained as a cream solid (42.6 mg, 0.0870 mmol, 70%).

From organoboronate ester: The title compound was synthesised from 7-iodo-7-deazaadenosine (19.6 mg, 0.0500 mmol) and 4-([4-methylsulfanylphenyl]ethynyl)phenylboronic acid neopentyl glycol ester (20.2 mg, 0.600 mmol, 1.2 eq). The product was obtained as a white solid (17.3 mg, 0.0350 mmol, 71%).

^{1}H NMR (400 MHz, DMSO-d_6) δ 8.16 (s, 1H, C2-H), 7.62 (d, J = 8.1 Hz, 2H, Ar-H), 7.53 (d, J = 8.1 Hz, 2H, Ar-H), 7.49 (s, 1H, C6-H), 7.45 (d, J = 8.3 Hz, 2H, Ar-H), 7.27 (d, J = 8.3 Hz, 2H, Ar-H), 6.11 (d, J = 6.3 Hz, 1H, C1'-H), 4.67 (dd, J = 6.2, 5.3 Hz, 1H, C2'-H), 4.31 (dd, J = 5.3, 3.0 Hz, 1H, C3'-H), 4.13 (app q, J = 3.0 Hz, 1H, C4'-H), 3.87 (dd, J = 12.4, 2.6 Hz, 1H, C5'-Ha), 3.75 (dd, J = 12.4, 2.9 Hz, 1H, C5'-Hb); ^{13}C NMR (126 MHz, DMSO-d_6, selected peaks from HSQC) δ 152.2, 132.2, 122.1, 128.9, 132.2, 126.1, 87.5, 73.9, 71.1, 85.7, 62.1, 62.2; ESI HRMS [MH]$^+$ 489.1587 (Calcd. For C$_{26}$H$_{25}$N$_4$O$_4$S 489.1581).

4-Amino-5-(4-[{3-thienyl}ethynyl]phenyl)-7-(β-D-ribofuranosyl)-7H-pyrrolo[2,3-d]pyrimidine 20d
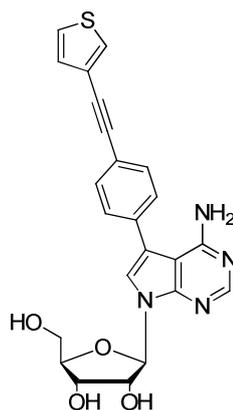
From organoboronate ester: The title compound was synthesised from 7-iodo-7-deazaadenosine (19.6 mg, 0.050 mmol) and 4-(3-thienylethynyl)phenylboronic acid neopentyl glycol ester (17.7 mg, 0.600 mmol, 1.2 eq). The product was obtained as a white solid (17.3 mg, 0.0385 mmol, 77%). ^{1}H NMR (400 MHz, CD$_3$OD) δ 8.15 (s, 1H, C2-H), 7.65 (dd, J = 3.0, 1.2 Hz, 1H, Ar-H), 7.61 (d, J = 8.5 Hz, 2H, Ar-H), 7.53 (d, J = 8.5 Hz, 2H, Ar-H), 7.49 (s, 1H, C6-H), 7.46 (dd, J = 5.0, 3.0 Hz, 1H, Ar-H), 7.21 (dd, J = 5.0, 1.2 Hz, 1H, Ar-H), 6.11 (d, J = 6.3 Hz, 1H, C1'-H), 4.67 (dd, J = 6.3, 5.4 Hz, C2'-H), 4.31 (dd, J = 5.4, 2.9 Hz, 1H, C3'-H), 4.13 (app dd, J = 5.5, 2.9 Hz, 1H, C4'-H), 3.87 (dd, J = 12.4 2.7 Hz, 1H, C5'-Ha), 3.75 (dd, J = 12.4, 2.9 Hz, 1H, C5'-Hb); ^{13}C NMR (126 MHz, CD$_3$OD) 159.0, 152.4, 151.6, 135.6, 133.1, 130.7, 129.93, 129.90, 126.9, 123.7, 123.6, 123.5, 118.1, 91.0, 89.2, 87.3, 86.3, 75.6, 72.5, 63.4, 49.4. ESI HRMS [MH]$^+$ 449.1264 (Calcd. for C$_{23}$H$_{21}$N$_4$O$_4$S 449.1278).

4-Amino-5-(4-[{4-trifluoromethylphenyl}ethynyl]phenyl)-7-(β-D-ribofuranosyl)-7H-pyrrolo[2,3-d]pyrimidine 20e
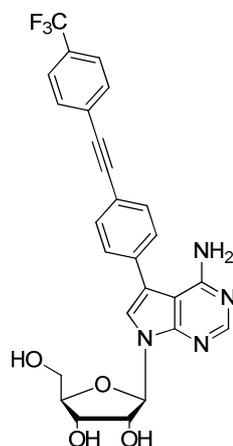
From organoboronate ester: The title compound was synthesised from 7-iodo-7-deazaadenosine   (19.6   mg,   0.050   mmol)   and   4-([4-trifluoromethylphenyl]ethynyl)phenylboronic acid neopentyl glycol ester (21.4 mg, 0.600 mmol, 1.2 eq). The product was obtained as a white solid (17.4 mg, 0.0341 mmol, 68%). ^{1}H NMR (400 MHz, CD$_3$OD) δ 8.14 (s, 1H, C2-H), 7.70 (d, J = 8.8 Hz, 2H, Ar-H), 7.68 (d, J = 8.8 Hz, 2H, Ar-H), 7.65 (d, J = 8.5 Hz, 2H, Ar-H), 7.55 (d, J = 8.5 Hz, 2H, Ar-H), 7.50 (s, 1H, C6-H), 6.09 (d, J = 6.3 Hz, 1H, C1'-H), 4.64 (dd, J = 6.3, 5.3 Hz, 1H, C2'-H), 4.28 (dd, J = 5.3, 2.9 Hz, 1H, C3'-H), 4.10 (app q, J = 2.8 Hz, 1H, C4'-H), 3.84 (dd, J = 12.4, 2.7 Hz, 1H, C5'-Ha), 3.73 (dd, J = 12.4, 2.9 Hz, 1H, C5'-Hb); ^{13}C NMR (101 MHz, CD$_3$OD) 158.94, 152.37, 151.58, 136.43, 133.43, 133.06, 130.87, 129.9 (q, J = 260 Hz, central two peaks of quartet observed), 129.97, 126.83, 126.5 (q, J = 3.8 Hz), 123.93, 122.63, 117.92, 92.50, 90.99, 89.41, 87.28, 75.62, 72.50, 69.57, 63.42; ESI HRMS [MH]$^+$ 511.1588 (Calcd. for C$_{26}$H$_{22}$F$_3$N$_4$O$_4$ 511.1588).

4-Amino-5-(4-[phenylethynyl]phenyl)-7-(2-deoxy-β-D-ribofuranosyl)-7H-pyrrolo[2,3-d]pyrimidine 35
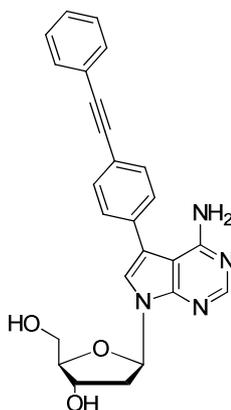
From organoboronate ester: The title compound was synthesised from 7-iodo-7-deaza-2'-deoxyadenosine (188 mg, 0.500 mmol) and 4-(phenylethynyl)phenylboronic acid neopentyl glycol ester (174 mg, 0.600 mmol, 1.2 eq). The product was isolated as a white solid (213 mg, 0.500 mmol, quant.). ^{1}H NMR (400 MHz, CD$_3$OD) δ 8.16 (s, 1H, C2-H), 7.64 (d, J = 8.5 Hz, 2H, Ar-H), 7.57 − 7.48 (m, 5H, Ar-H), 7.44 − 7.34 (m, 3H, Ar-H), 6.60 (dd, J = 8.2, 6.0 Hz,

1H, C1'-H), 4.55 (app dt, J = 5.5, 2.7 Hz, 1H, C3'-H), 4.03 (app q, J = 3.3 Hz 1H, C4'-H), 3.82 (dd, J = 12.1, 3.3 Hz, 1H, C5'-Ha), 3.74 (dd, J = 12.1, 3.6 Hz, 1H, C5'-Hb), 2.72 (ddd, J = 13.4, 8.2, 6.0 Hz, 1H, C2'-Ha), 2.36 (ddd, J = 13.4, 6.0, 2.7 Hz, 1H, C2'-Hb); ^{13}C NMR (126 MHz, CD$_3$OD) δ 158.8, 152.2, 151.3, 135.8, 133.2, 132.5, 129.9, 129.58, 129.55, 124.5, 123.5, 123.1, 118.2, 102.5, 91.0, 89.9, 89.1, 86.5, 73.0, 63.7, 41.5; ESI HRMS [MH]$^+$ 427.1746 (Calcd. For C$_{25}$H$_{23}$N$_4$O$_3$ 427.1765).

4-{[(Dimethylamino)methylidene]amino}-5-(4-[phenylethynyl]phenyl)-7-(5-*O*-[{4,4'-di methoxytriphenyl}methyl]-(2-deoxy-β-D-ribofuranosyl)-7H-pyrrolo[2,3-d]pyrimidine 36
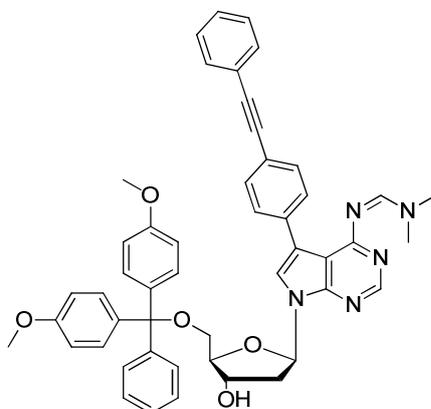
Based on the procedure reported by Seela *et al.*[128]

Compound **35** (171 mg, 0.400 mmol) was stirred with DMF dimethylacetal (640 µl) in MeOH (5.7 ml) under a N$_2$ atmosphere for 2 h at 40 °C. The volatiles were then removed under vacuum. Pyridine (2 ml) was then added, followed by DMTrCl (203 mg, 0.600 mmol, 1.5 eq). The reaction as then stirred at 50 °C for 1 h. The reaction was then quenched with 5% aq. NaHCO$_3$ (7 ml) and extracted with CH$_2$Cl$_2$ (3 x 20 ml). The organic layers were dried over Na$_2$SO$_4$ and the volatiles removed *in vacuo*. The crude product was purified by column chromatography on silica gel, eluting with MeOH/Et$_3$N/CH$_2$Cl$_2$ (0:0.5:99.5 → 5:0.5:94.5) to give the product as a white solid (229.2 mg, 73%). ^{1}H NMR (500 MHz, CD$_2$Cl$_2$) δ 8.80 (s, 1H, C2-H), 8.45 (s, 1H, C6-H), 7.69 (d, J = 8.4 Hz, 2H, Ar-H), 7.60 (dd, J = 7.8, 1.6 Hz, 2H, Ar-H), 7.50 – 7.38 (m, 8H, Ar-H), 7.35 (d, J = 8.9 Hz, 4H, Ar-H), 7.32 – 7.22 (m, 3H, Ar-H), 6.85 (app t, J = 6.9 Hz, 1H, C1'-H), 6.83 – 6.78 (m, 4H, Ar-H), 4.74 (app dt, J = 6.4, 3.3 Hz, 1H, C3'-H), 4.14 (app q, J = 3.9 Hz, 1H, C4'-H), 3.782 (s, 3H, OCH$_3$), 3.778 (s, 3H, OCH$_3$), 3.44 (dd, J =

10.2, 4.2 Hz, 1H, C5'-Ha), 3.34 (dd, *J* = 10.2, 4.2 Hz, 1H, C5'-Hb), 3.16 (s, 3H, NCH$_3$), 3.06 (s, 3H, NCH$_3$), 2.76 (ddd, J = 13.5, 7.5, 6.2 Hz, 1H, C2'-Ha), 2.53 (ddd, J = 13.5, 6.3, 3.5 Hz, 1H, C2'-Hb), 2.30 (br s, 1H); ^{13}C NMR (126 MHz, CD$_2$Cl$_2$) δ 161.6, 159.0, 156.7, 153.3, 152.0, 145.2, 136.2, 136.0, 135.4, 131.8, 130.9, 130.4, 130.2, 128.8, 128.6, 128.5, 128.3, 127.2, 123.8, 121.7, 120.8, 118.7, 113.47, 113.46, 109.3, 90.2, 89.5, 86.9, 85.9, 83.4, 73.3, 64.5, 55.6, 41.2, 41.1, 35.4; ESI HRMS [MH]$^+$ 783.3420 (Calcd. for C$_{49}$H$_{45}$N$_5$O$_5$ 783.3415).

4-{[(Dimethylamino)methylidene]amino}-5-(4-[phenylethynyl]phenyl)-7-(3-*O*-[2-cyanoethyl diisopropylphosphoramidite]-5-*O*-[{4,4'-dimethoxytriphenyl}methyl]-(2-deoxy-β-D-ribofuranosyl)-7H-pyrrolo[2,3-d]pyrimidine[128] 37
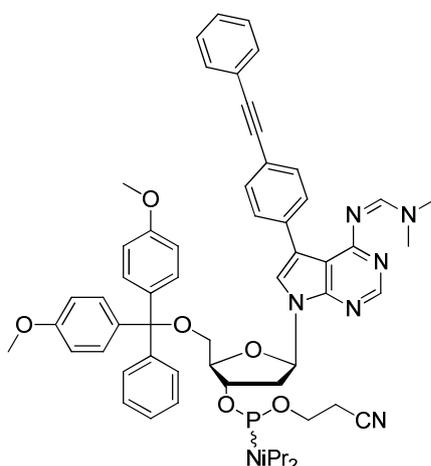
Based on the procedure reported by Seela *et al.*[128]

Compound **36** (220 mg, 0.281 mmol) was dissolved in THF (1.2 ml). Freshly distilled ^{i}Pr$_2$EtN (166 µl, 0.927 mmol, 3.3 eq) was added, followed by 2-cyanoethyl-*N,N*-diisopropylchlorophosphoramidite (78 µl, 0.35 mmol, 1.25 eq). The reaction was stirred at r.t. for 2 h. The mixture was then filtered, diluted with EtOAc (15 ml) and washed with cold aq. Na$_2$CO$_3$ (1 M, 2 x 5 ml) and water (5 ml). The organic layer was dried over Na$_2$SO$_4$ and the volatiles removed *in vacuo*. The crude product was purified by column chromatography on silica gel, eluting with acetone/Pet Ether 40-60 (3:7) containing 0.5% Et$_3$N. The product was isolated as a low melting solid which contained an unknown impurity (impurity present in 2-cyanoethyl-*N,N*-diisopropylchlorophosphoramidite starting material) Yield = 148 mg. Selected NMR peaks given below. ^{1}H NMR (400 MHz, CD$_2$Cl$_2$) δ 8.76 (s), 8.41 (s), 8.40 (s), 7.71 (s), 7.64 (d, *J* = 8.2 Hz, Ar-H), 7.59-7.50 (m, Ar-H), 7.48-7.32

(m, Ar-H), 7.35-7.25 (m, Ar-H), 7.27-7.17 (m, Ar-H), 6.80 (t, *J* = 7.0 Hz, C1'-H), 6.79-6.69 (m, Ar-H), 6.11 (s), 4.91-4.67 (m), 4.29-4.19 (m), 3.73 (s), 3.73 (s), 3.72 (s), 3.39-3.28 (m), 3.11 (s, NCH$_3$), 3.02 (s, NCH$_3$), 2.64 (t, *J* = 6.3 Hz, C2'-Ha), 2.49 (t, *J* = 6.3 Hz, C2'-Hb), 1.24-1.15 (m); ^{13}C NMR (126 MHz, CD$_2$Cl$_2$) δ 156.3, 151.9, 129.8, 130.4, 130.3, 130.0, 40.8, 35.1; ^{31}P NMR (212 MHz, CD$_2$Cl$_2$) δ 149.06, 149.05; ESI HRMS [MH]$^+$ 984.4541 (Calcd. for C$_{58}$H$_{63}$N$_7$O$_6$P 984.4572).

4-Amino-5-(4-[phenylethynyl]phenyl)-7-(5-triphosphate-β-D-ribofuranosyl)-7H-pyrrolo[2,3-d]pyrimidine 43
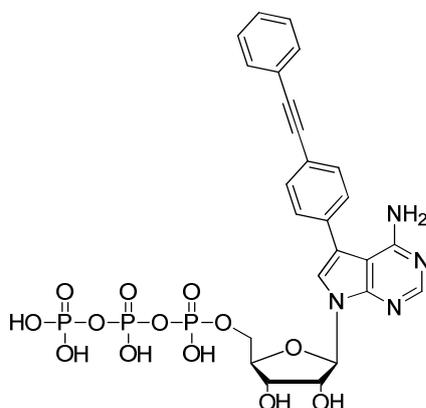
Preparation of tri-*n*-butylammonium pyrophosphate

Tetrasodium diphosphate was treated with Dowex 50WX8 (H$^+$ form) to give tri-*n*-butylammonium pyrophosphate [*Can. J. Chem.* (1964) 42:599-604]. Tetrasodium diphosphate decahydrate (2.23 g, 5 mmol) was dissolved in water (50 ml) with sonication and applied to the Dowex column. The column was washed with water and the eluate was dropped directly into a cooled (ice water) and stirred solution of ^{n}Bu$_3$N (2.38 ml, 10 mmol) in ethanol (20 ml). The column was washed until the pH of the eluate increased to 5.0 (*ca.* 70 ml of water). The ethanol/water solution was evaporated to dryness. The tri-*n*-butylammonium pyrophosphate product was co-evaporated with 3x50 ml of methanol to remove water (< 40˚C to prevent decomposition). The product was then dried under high vacuum overnight. The residue was dissolved in anhydrous DMF (10 ml) and used within 24 h.

Ludwig conditions[200]

The nucleoside (88.4 mg, 0.2 mmol) was dissolved in dry trimethylphosphate (1 ml) under a N$_2$ atmosphere and cooled to 0 °C. Phosphoryl chloride (55.9 μL, 0.6 mmol, 3 eq) was added and the mixture stirred for 3 h at 0 °C. Tri-n-butylammonium pyrophosphate in anhydrous DMF (2 mmol, 3 ml) was then added, followed by ^{n}Bu$_3$N (0.6 ml). After stirring for 30 min, the reaction was quenched by addition of aqueous triethylammonium hydrogen carbonate buffer (20 ml, pH 7.5, 1 M). After stirring at rt for 30 min, the solvent was removed *in vacuo* (< 40˚C) to yield an off-white solid.

The product was resuspended in *ca.* 10 ml TEAB buffer (pH 7.5, 0.05 M) and filtered twice through 0.45 μm pore syringe filter cartridges. The product was purified using fast protein liquid chromatography (FPLC) on a DEAE-Sephadex (A25-Chloride form, 20 ml) column and TEAB buffer (0.05 to 1.2 M linear gradient, pH 7.5). The product was loaded onto the column through the pump, and washed through with 0.05 M TEAB buffer until the absorbance of the eluent stopped changing.

The relevant fractions are pooled and concentrated under reduced pressure to *ca.* 10 ml. Methanol (50 ml) was added and the compound concentrated to *ca.* 10 ml. This procedure was repeated three times. The solution was then cooled to 0 °C and evaporated to dryness under high vacuum.

The compound was then redissolved in a minimum amount of methanol, and a saturated solution of sodium iodide in acetone was added dropwise (0 °C, kept in dark), leading to precipitation of a solid. The solid was filtered (sintered glass filter) and washed with acetone (10 ml), Et$_2$O (10 ml) and allowed to air dry (alternatively, the solid could be isolated by centrifugation). The solid was washed through the filter with water (10 ml) and dried overnight under high vacuum. Mass spectrometry (ESI⁻) of the crude product found the mass ion corresponding to the triphosphate product, as well as diphoshpate and monophosphate products. Attempts to isolate the product by semi-preparative RP-HPLC failed.

Huang conditions[140]

Salicyl chlorophosphite (40.5 mg, 0.2 mmol, 2 eq) dissolved in dry DMF (0.3 ml). ^{n}Bu3N (0.38 ml) was added to tri-n-butylammonium pyrophosphate in DMF (0.3 ml) in a separate

flask under a N$_2$ atmosphere. The pyrophosphate mixture was then added to the salicyl chlorophosphite and the mixture was stirred at rt for 30 min, and then cooled to 0 °C. The mixture was then added to a Schlenk tube containing the nucleoside (44.2 mg, 0.1 mmol). The nucleoside phosphate mixture was stirred at 0 °C for 5 h and then allowed to reach rt. A solution of iodine (1.25 ml, 3% *w/v* in 9:1 pyridine/water) was added dropwise to the reaction mixture (until the reaction maintains the iodine colour). The reaction was then stirred for 15 min. Water (2.5 ml) was then added to hydrolyse the product, and the mixture stirred for 1.5 h. The reaction mixture was then transferred to a 50 ml falcon tube. NaCl solution (0.7 ml, 3 M) was then added, followed by ice-cold EtOH (19 ml) to precipitate the product. The precipitate was isolated by centrifugation (5000 g, 10 min, 5 °C). The supernatant was decanted and the precipitate allowed to air dry. A mixture of products was identified by HPLC, but could not be separated. ^{31}P NMR suggested the presence of a mixture of triphosphate products.

UV-Vis spectroscopy

Extinction coefficients for nucleosides were measured in solvents as described by taking spectra over five concentrations (between 1×10^{-4} M and 1×10^{-6} M). Solutions of nucleosides for UV-Vis spectra in DCM, EtOAc and water were made from a 2.5×10^{-3} M stock in DMSO. UV-Vis spectra of oligonucleotides were measured in EB buffer at *ca.* 5 μM. All measurements were made in triplicate.

Fluorescence quantum yields

Fluorescence quantum yields were measured using an integrating sphere in a Horiba Fluoromax 3 fluorimeter using solutions with an absorbance of *ca.* 0.1 (for the nucleosides), or a molecular concentration of 2.5 μM (for the oligonucleotides), according to the manufacturer's instructions. The oligonucleotide quantum yields were measured in a buffered solution at pH 7.0 (50 mM sodium phosphate, 0.1 M NaCl).

Melting temperatures

The oligonucleotide and the complementary strand were dissolved in a buffered solution at pH 7.0 (50 mM sodium phosphate, 0.1 M NaCl) to give 850 μl of a solution containing the duplex (2.5 μM). The solutions were then degassed for 10 min with stirring, and layered with 150 μl of mineral oil. Absorbance vs temperature curves were measured at 260 nm ramping from 20 °C to 90 °C and then from 90 °C to 20 °C, at 1 °C/min with stirring at *ca.* 800 rpm. The raw data was differentiated over 6-point steps, and the first derivatives fitted to Gaussian curves using OriginLab to find the melting temperatures. All measurements were repeated in triplicate.

Fluorescence titration measurements

850 μL of a solution of the oligonucleotide (2.5 μM) in pH 7.0 buffer (50 mM sodium phosphate, 0.1 M NaCl) was prepared. The excitation spectrum (200-600 nm, emission measured at 450 nm) was measured to find the excitation maximum. The emission spectrum was then measured (365-600 nm, λ_{ex} = 322 nm, slit width = 1 nm) at 20 °C. Several additions of the complementary strand were then made, with thorough mixing of the sample after each addition (typically each addition was 0.25 eq with respect to the

oligonucleotide, to give a final ratio of 1:2 oligo:complementary strand). The emission spectrum was remeasured (as above) after each addition.

Fluorescence quenching experiments

Fluorescence quenching experiments were carried out using 2.5 µM solutions of the oligonucleotides or duplexes in pH 7.0 buffer (as above) at 20 °C or 60 °C. Acrylamide (Acrylagel, gas stabilised 30% w/v aqueous solution) was added portion-wise to give solutions with the specified acrylamide concentrations. Fluorescence emission spectra were measured as above after each addition.

Effect of pH on fluorescence

The pH dependence of fluorescence spectra was investigated using 2.5 µM solutions of the oligonucleotides or duplexes in buffered solution at 20 °C. Sodium phosphate buffer was used for pH 6, 7, 8 and 9. Sodium citrate buffer was used for pH 3, 4, and 5. Sodium borate buffer was used for pH 10 and 11. pK_a values were determined by fitting the data to a sigmoidal Boltzmann function using OriginLab according to the equation:

$$y = \frac{A_1 - A_2}{1 + e^{(x - x_0)/dx}} + A_2$$

Where the pK_a is equal to $(A_1 + A_2)/2$.

Circular dichroism

Circular dichroism spectra were measured at 200 µM (base concentration, 7.69 µM duplex concentration) in pH 7.0 buffer (as above) in a cuvette with a path length of 0.1 cm. The spectra were measured over 200-320 nm, scanning at 50 nm/min, at 10 °C and 75 °C.

Fluorescence Lifetime Measurements[201]

Samples for fluorescence lifetime measurement were prepared so as to have an absorbance of *ca.* 0.1 at the absorbance maximum. The fluorescence lifetimes were measured by time-correlated single photon counting (TCSPC). The excitation source used was the 3rd harmonic (300 nm) of a mode-locked (900 nm), cavity dumped (APE Pulse switch) Ti:sapphire laser (Coherent MIRA 900), pumped by the 2nd harmonic (532 nm) of a

CW Nd:YAG laser (Coherent Verdi V6). The pulse characteristics were as follows: a temporal full width at half maximum (FWHM) of ~150 fs, average power 0.3 mW at a repetition rate of 4 MHz. The fluorescence emission was collected at right angles to the excitation source, with the emission wavelength selected by a monochromator (Jobin Yvon TRIAX 190) and detected by a cooled photomultiplier tube module (IBH TBX-04). The detector was linked to a time-to-amplitude converter (Ortec 567) and a pulse height analyser, PHA, (E.G. & G. Trump Card) controlled by the software Maestro (ver 5.10). Fluorescence decays were recorded to a minimum of 10,000 counts in the peak channel of the PHA with a record length of 1000 channels. The band pass of the monochromator was adjusted to give a signal count rate of 5-20 KHz. The instrument response function (IRF) was measured using a dilute LUDOX® colloidal silica suspension as the scattering sample, giving an IRF of ~200 ps FWHM. Iterative reconvolution of the IRF with one or two decay functions and non-linear least-squares analysis were used to analyse the data in Microsoft Excel using the solver function. The quality of the fit was judged by the calculated value of the reduced χ^2 and Durbin-Watson parameters and visual inspection of both the residual and the autocorrelated residuals.

General procedure for direct arylation of 2′-deoxyadenosine at 60 °C

2′-Deoxyadenosine (251 mg, 1.0 mmol, 1.0 eq), CuI (571 mg, 3.0 mmol, 3 eq), caesium carbonate (815 mg, 2.5 mmol, 2.5 eq) and the palladium source (0.7-5 mol%) were added to a Schlenk tube under a N_2 atmosphere. DMF (5.3 ml) and the aryl iodide (2.0 mmol, 2.0 eq) were added and the reaction was stirred at 60 °C for either 24 or 48 h. The reaction was then cooled to room temperature, and HCl (5.3 ml, 1 M) was added to acidify the crude mixture. The pH was tested, and NaOH (1 M) was added until the pH reached 6.5. The reaction mixture was then extracted with EtOAc/^{i}PrOH (5 x 25 ml, 9:1 v/v), dried over $MgSO_4$ and the solvent removed *in vacuo*. The crude product was redissolved in MeOH/CH_2Cl_2 (1:1 v/v) and absorbed onto silica. The product was then purified by column chromatography on silica gel, eluting with MeOH/CH_2Cl_2 (2:98 → 1:9 *v/v*). The solvent was removed *in vacuo* to yield the pure product. Characterisation as previously reported.[86]

$Pd_2(dba-4-OMe,4'-CF_3)_3$

NaCl (133 mg, 2.27 mmol, 2.0 eq) was added to a solution of $PdCl_2$ (200 mg, 1.13 mmol) in methanol (10 ml) and stirred at room temperature under an inert atmosphere for 24 h. The solution was then filtered through a plug of cotton wool. The solution was warmed to 60 °C and dba-4-OMe,4′-CF_3 (1.16 g, 3.50 mmol, 3.1 eq) was added. The warm mixture was stirred vigorously, and methanol added until the ligand was fully dissolved (*ca.* 60 ml). Sodium acetate (556 mg, 6.78 mmol, 6.0 eq) was then added to the solution. The reaction mixture was removed from the heat and allowed to cool to ambient temperature (with no external cooling), and stirred for 2 h until a dark red/brown precipitate was formed. The precipitate was filtered and washed with methanol (2 x 5 ml), water (2 x 2 ml), and cold acetone (2 x 0.5 ml). The product was dried by suction to give the complex as a dark red/brown solid (848 mg, 97 %). Mp 163-168 °C; Anal. calcd. for $C_{57}H_{45}F_9O_6Pd_2$: C 56.59, H 3.75; found: C, 56.48; H, 3.67.

Pd(PVP) nanoparticles (*ca.* 2 nm)

Prepared according to Ellis *et al.*[184]

$PdCl_2$ (85.1 mg, 0.48 mmol) was stirred with HCl (aq., 0.20 M, 4.8 ml) in water (235 ml) and heated slightly until the $PdCl_2$ dissolved (indicating formation of $H_2[PdCl_4]$, 2.0 mM solution). Water (224 ml) and EtOH (336 ml) were added, followed by PVP (M_w = 29000,

1.067 g). The reaction mixture was heated to reflux for 3 h with vigorous stirring. During this time the reaction mixture formed a dark brown solution. The reaction mixture was allowed to cool to room temperature and the solvents removed *in vacuo*. The nanoparticles were dried overnight under high vacuum, and then ground to a fine powder. The particles were characterised by TEM.

Preparation of samples for TEM

The direct arylation reaction was carried out as described above. After 1.5 h, a 1 ml aliquot was removed and PVP (M_w = 29,000; 208 mg, 20 monomer equivalents with respect to Pd) was added and the mixture was stirred for 30 min. The volatiles were then removed *in vacuo*. The solid samples were suspended in ethanol and dropped onto a gold grid. The nanoparticles were characterised using a FEI Tecnai 12 Biotwin High Contrast Electron Microscope operated at 120 keV. The nanoparticles were measured using the PSA macro for ImageJ.

Appendix I – UV-Vis and Fluorescence Spectra

4-Amino-5-phenylethynyl-7-(β-D-ribofuranosyl)-7H-pyrrolo[2,3-d]pyrimidine 16
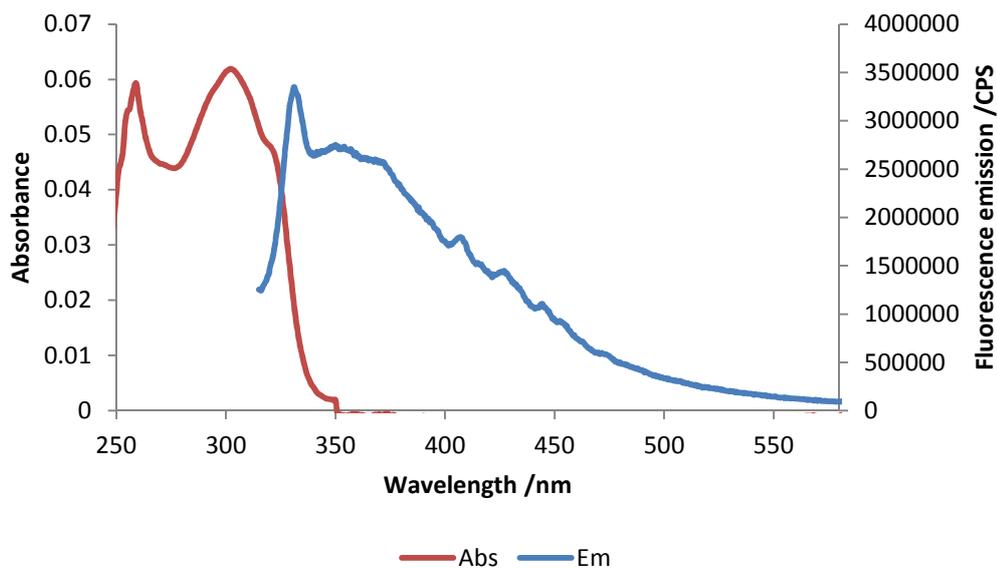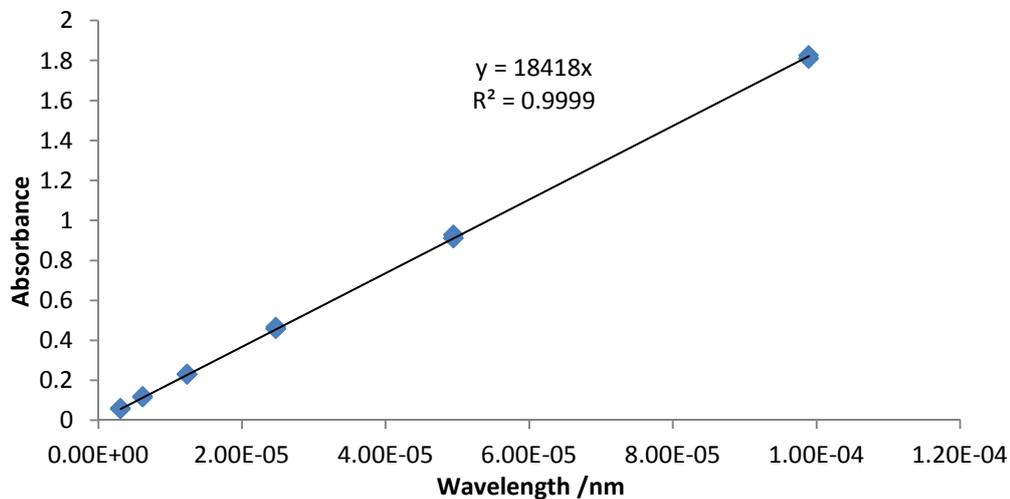
4-Amino-5-[(1-naphthyl)ethynyl]-7-(β-D-ribofuranosyl)-7H-pyrrolo[2,3-d]pyrimidine 17
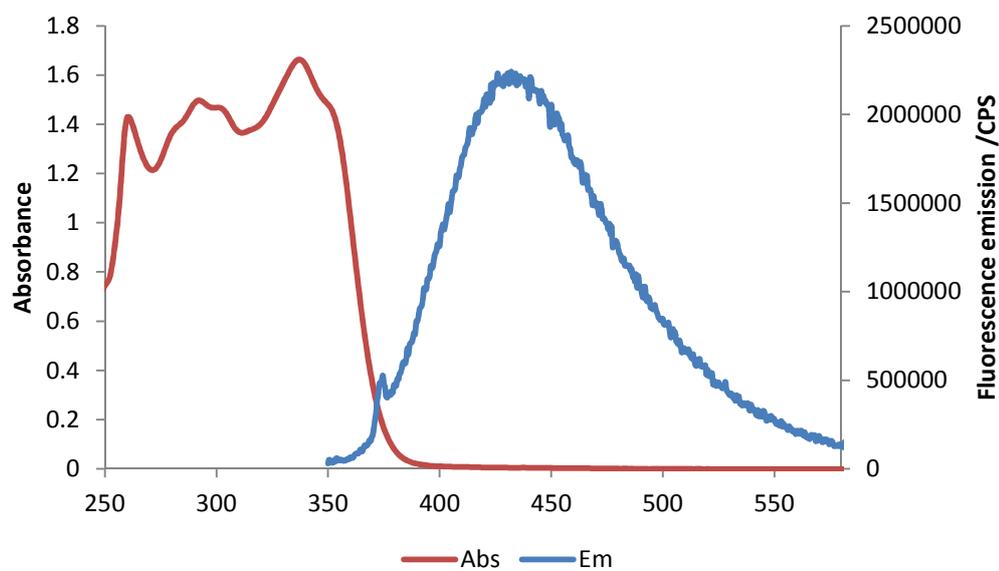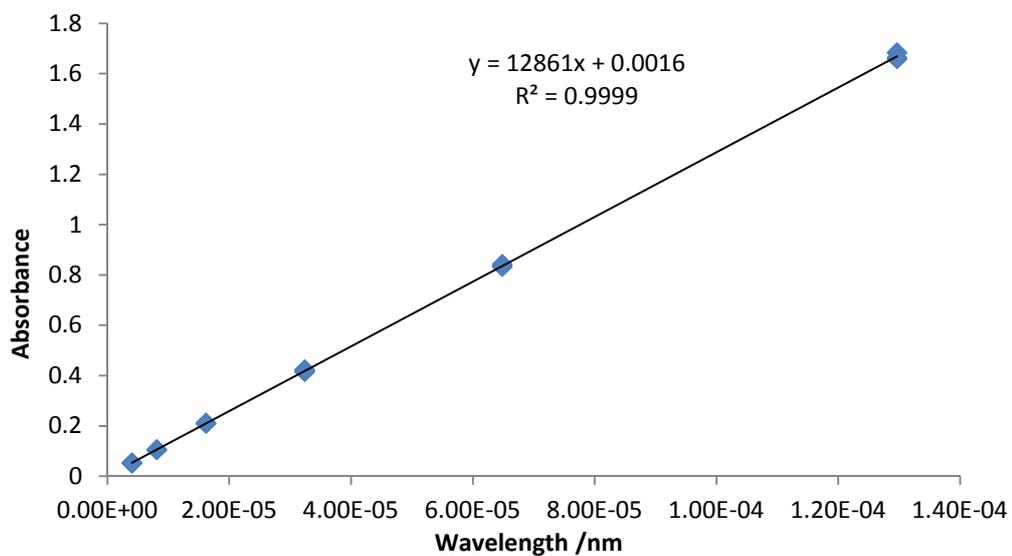
$y = 12861x + 0.0016$
$R^2 = 0.9999$

4-Amino-5-phenyl-7-(β-D-ribofuranosyl)-7H-pyrrolo[2,3-d]pyrimidine 18

$y = 12402x + 0.0012$
$R^2 = 0.9999$
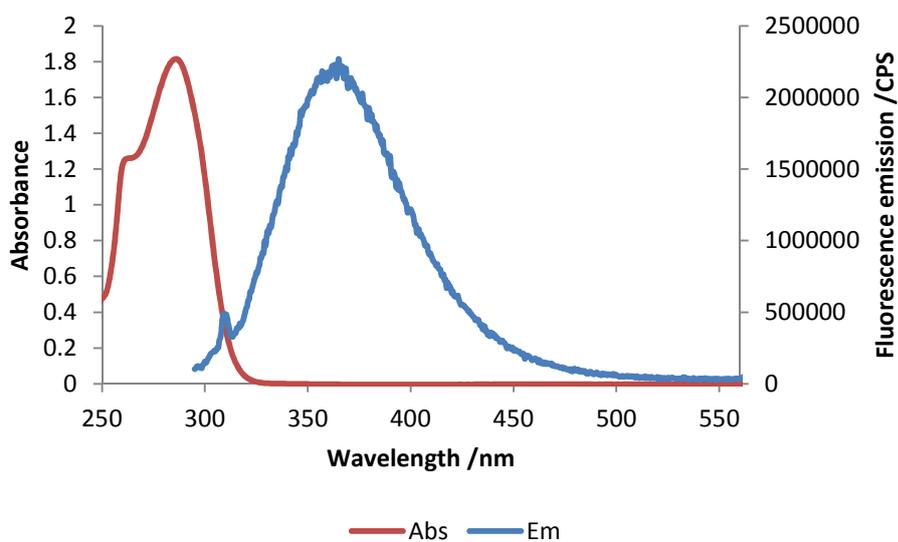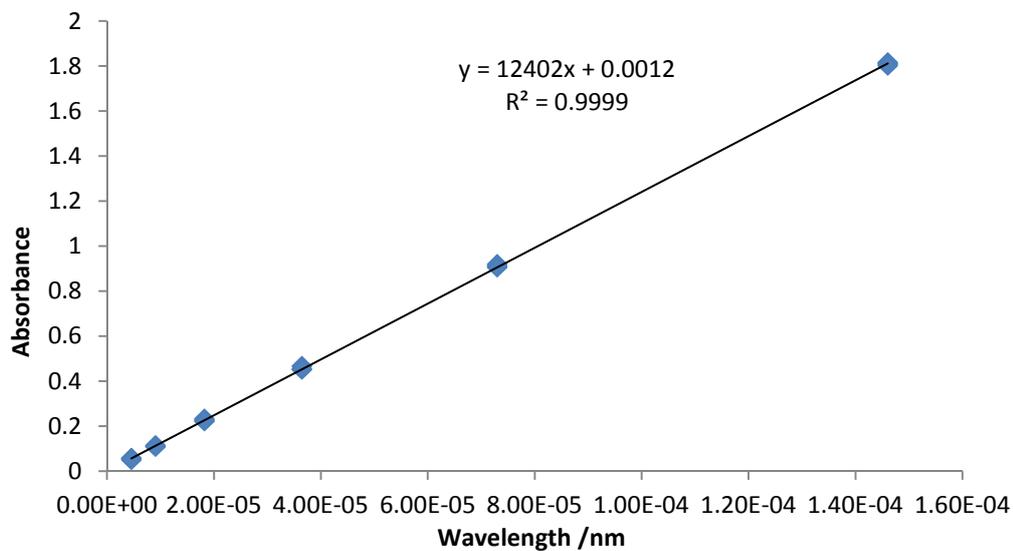
4-Amino-5-(4-[phenylethynyl]phenyl)-7-(β-D-ribofuranosyl)-7H-pyrrolo[2,3-d]pyrimidine 20a
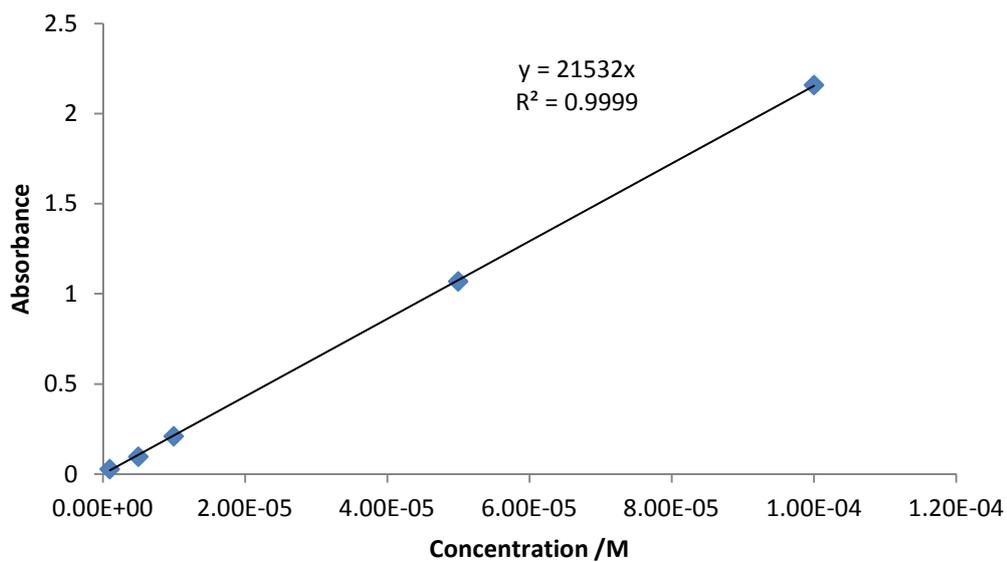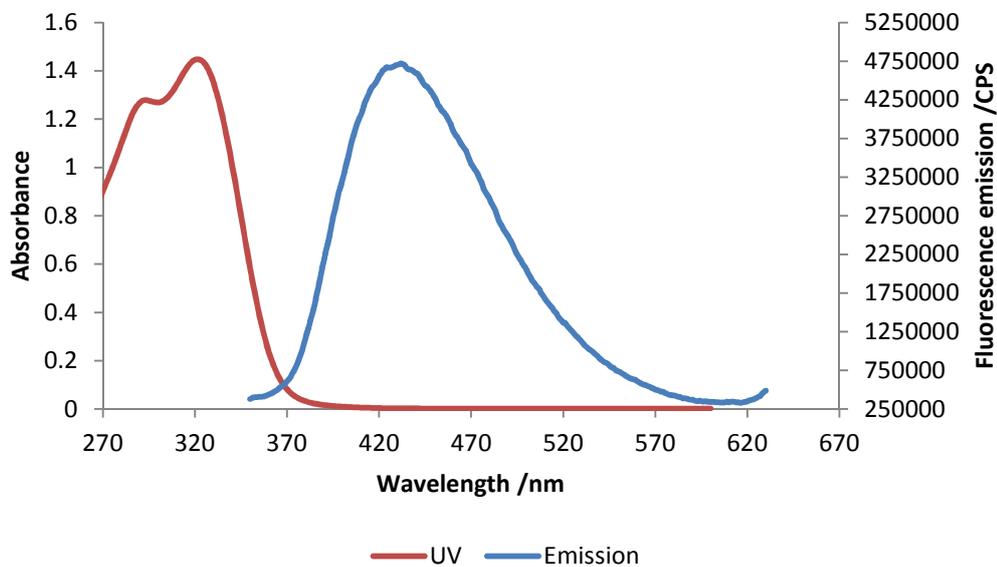
4-Amino-5-(4-[{4-Methoxyphenyl}ethynyl]phenyl)-7-(β-D-ribofuranosyl)-7H-pyrrolo[2,3-d]pyrimidine 20b

4-Amino-5-(4-[{4-Methylsulfanylphenyl}ethynyl]phenyl)-7-(β-D-ribofuranosyl)-7H-pyrrolo[2,3-d]pyrimidine 20c

y = 25359x
R² = 0.9994

4-Amino-5-(4-[{3-thienyl}ethynyl]phenyl)-7-(β-D-ribofuranosyl)-7H-pyrrolo[2,3-d]pyrimidine 20d

$y = 21444x$
$R^2 = 1$

4-Amino-5-(4-[{4-trifluoromethylphenyl}ethynyl]phenyl)-7-(β-D-ribofuranosyl)-7H-pyrrolo[2,3-d]pyrimidine 20e

$$y = 14861x$$
$$R^2 = 0.9999$$
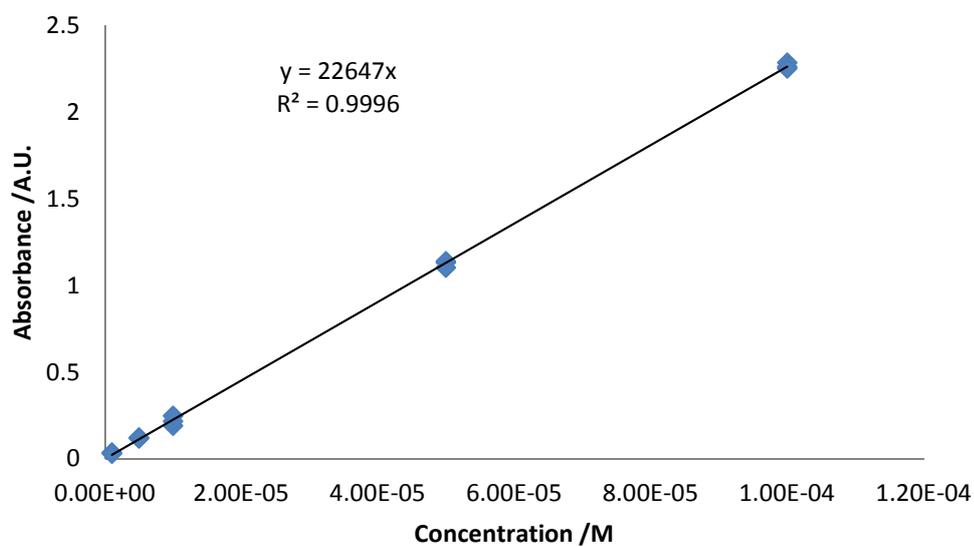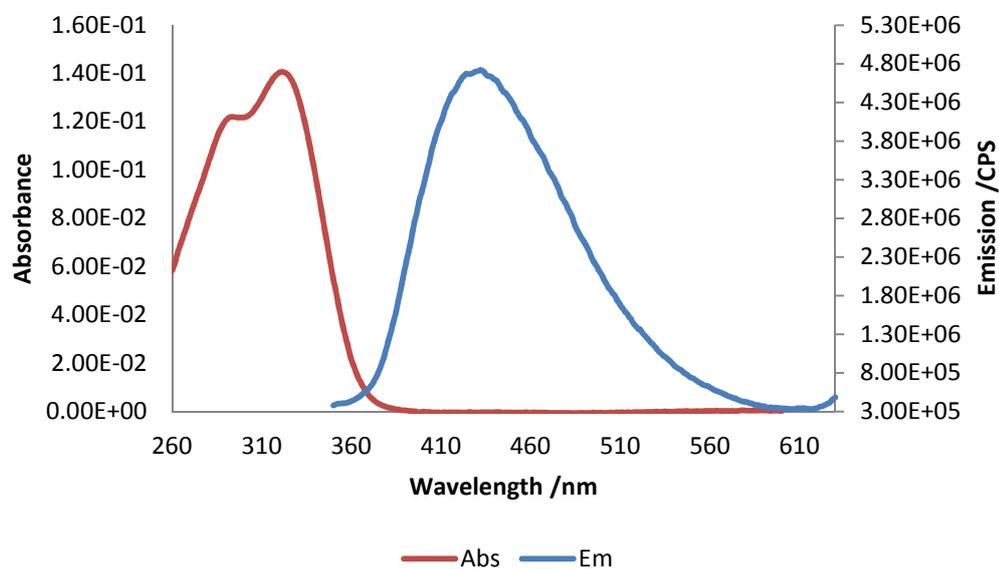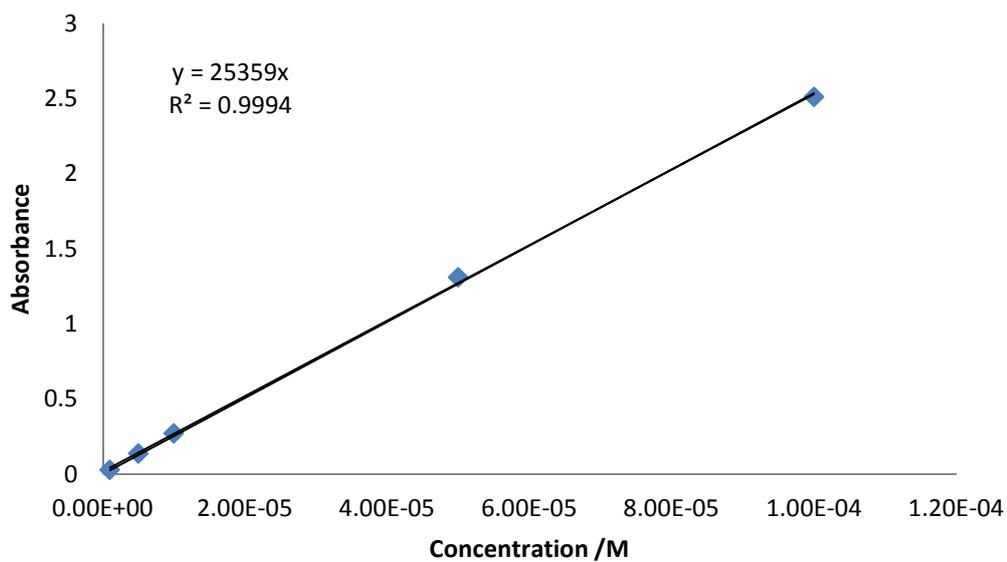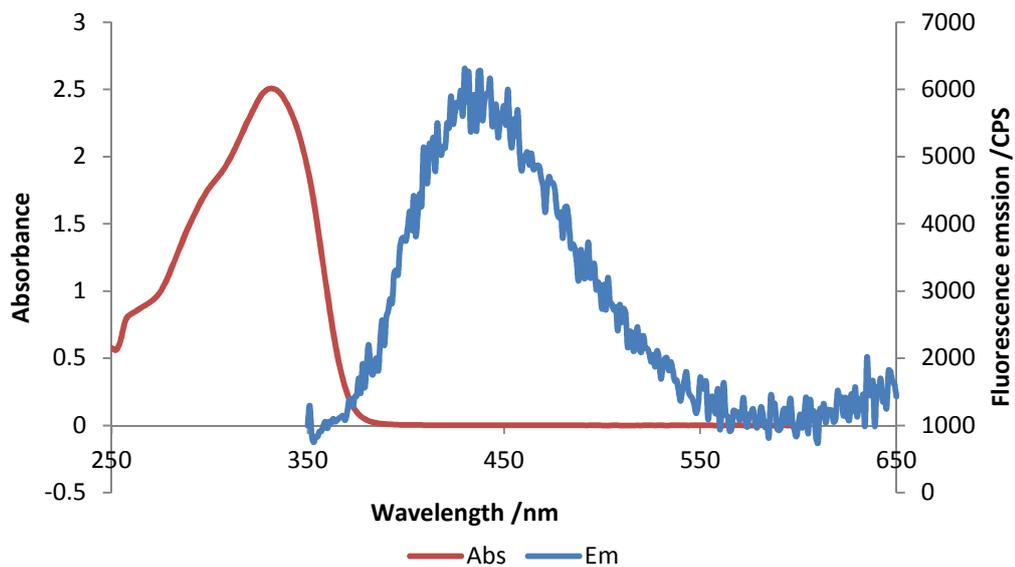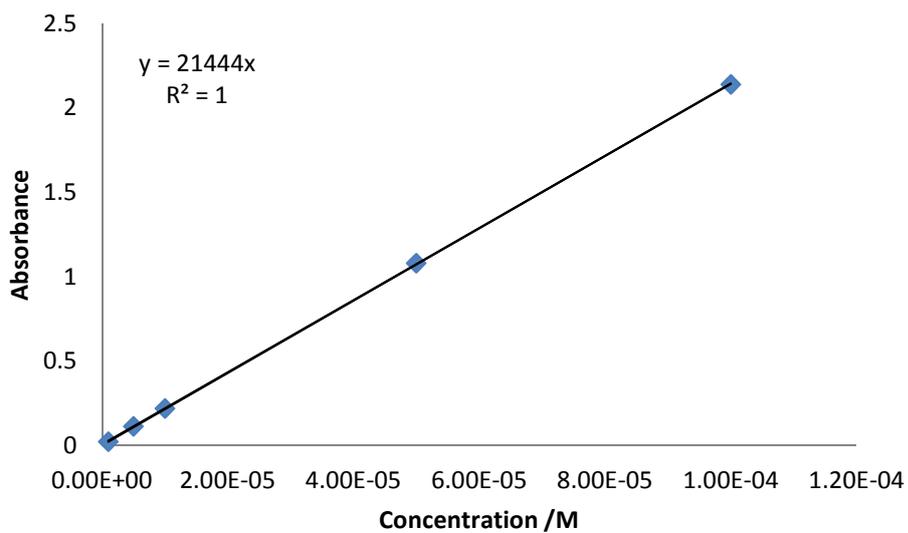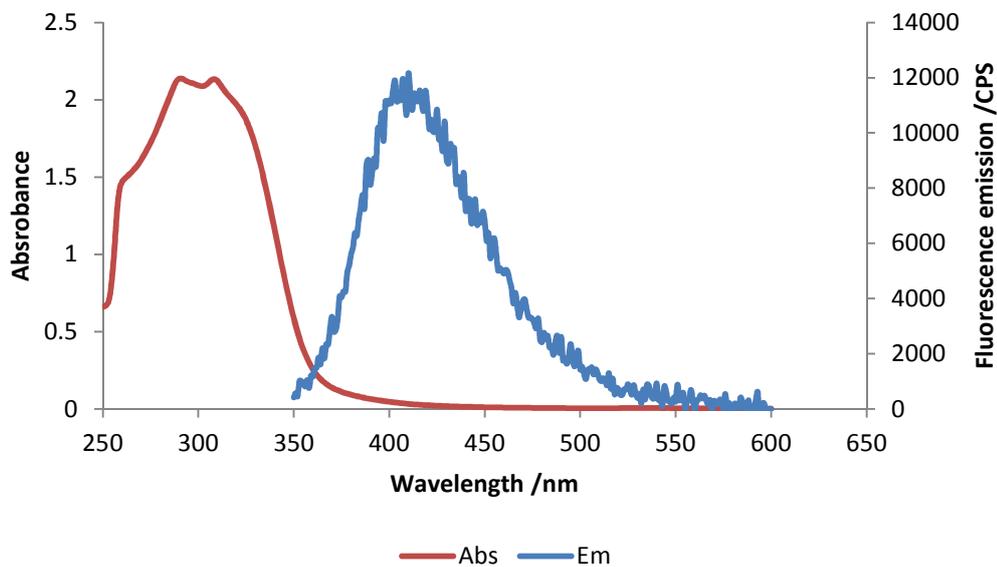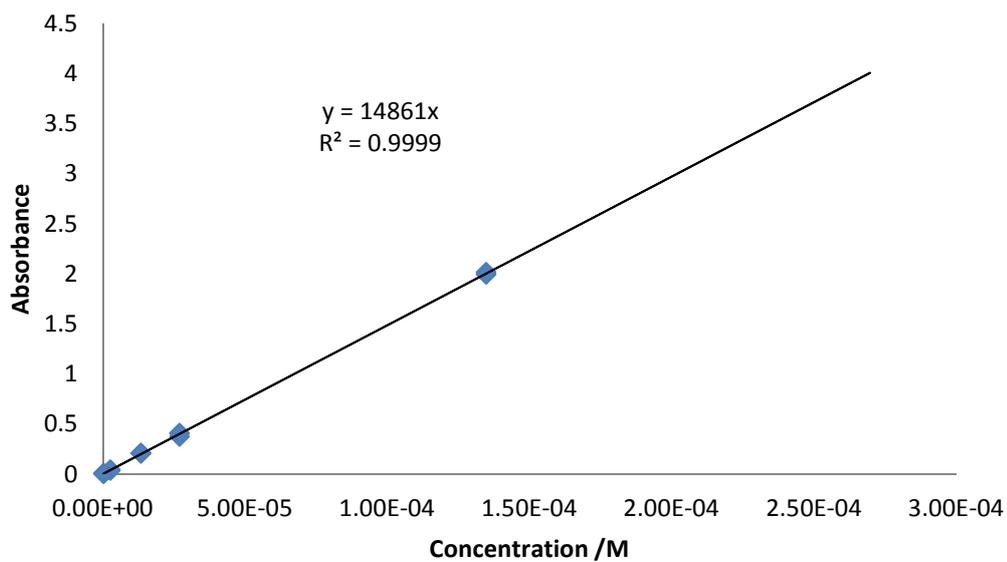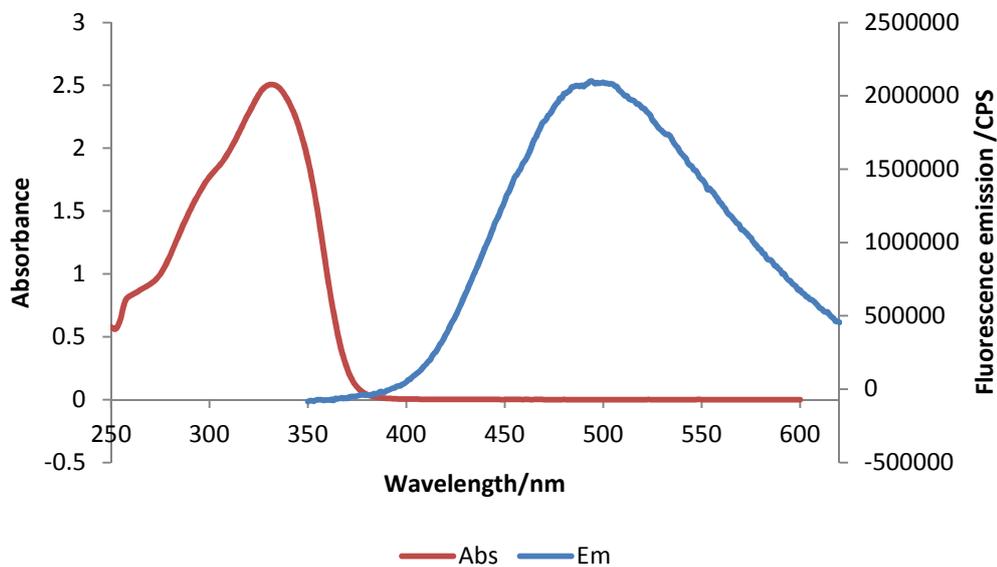
Fluorescence Lifetime Measurements

Compound	Emission Wavelength /nm	Fluorescence lifetime /ns	Durbin-Watson parameter
16	350	4.1 (87%), 0.5 (13%)	1.73
17	430	1.5 (60%), 0.5 (40%)	1.74
18	360	2.0 (37%), 0.2 (63%)	1.12
20a	430	2.2 (92%), 0.5 (8%)	1.83
20b	405	1.6 (90%), 0.5 (10%)	1.92
20c	430	2.0 (92%), 0.5 (8%)	1.86
20d	405	1.5 (88%), 0.5 (12%)	1.82
20e	490	2.1 (87%), 0.8 (12%)	1.82
35	430	2.3 (88%), 0.6 (12%)	1.92
35 (water)	430	0.4 (62%), 0.9 (38%)	1.83
35 (water)	500	0.4 (61%), 1.0 (39%)	1.83

Appendix II - Melting curves

Melting curves are shown as the change in UV absorbance at 260 nm against temperature in °C. The differentiated and fitted data is shown below the corresponding melting curves, where A (vertical axis) is the derivative and B (horizontal axis) is the temperature in °C.

ODN1-1T
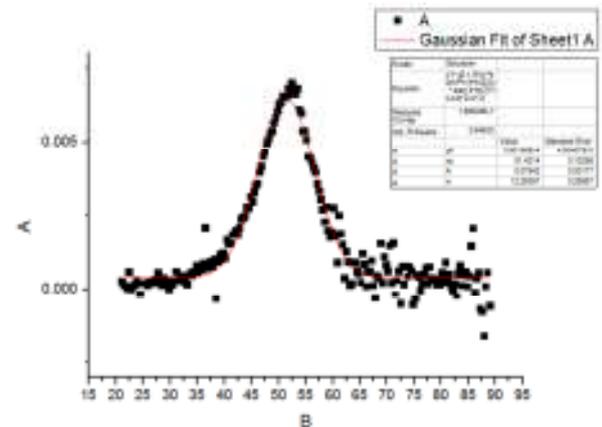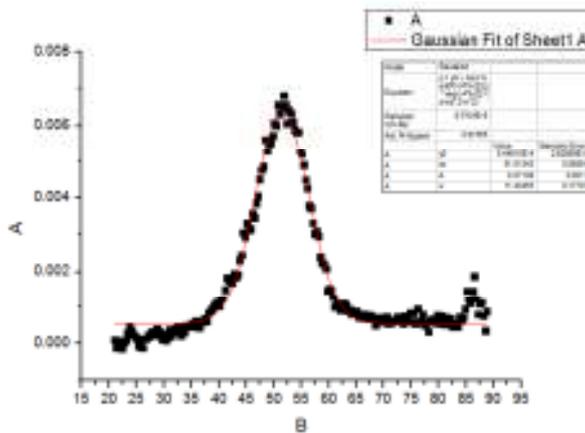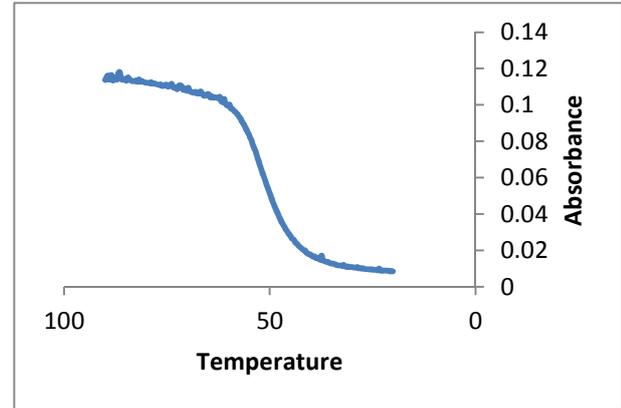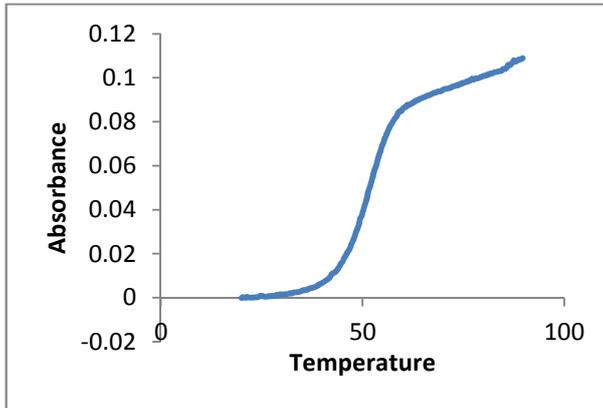
ODN1-1C
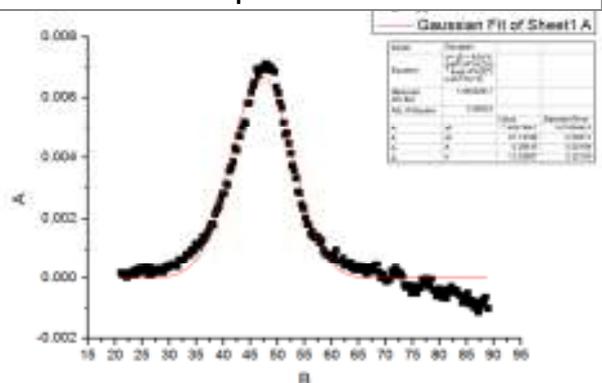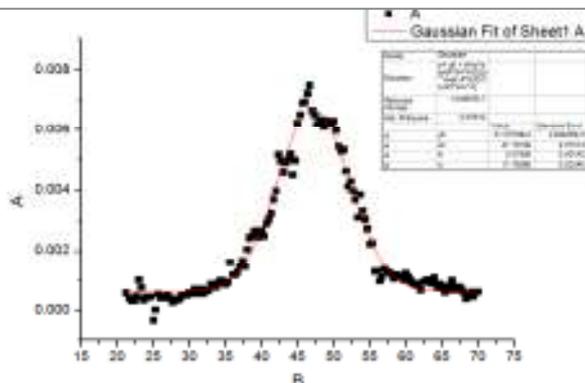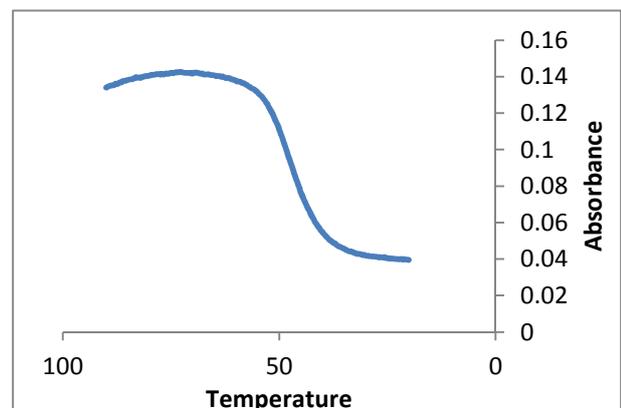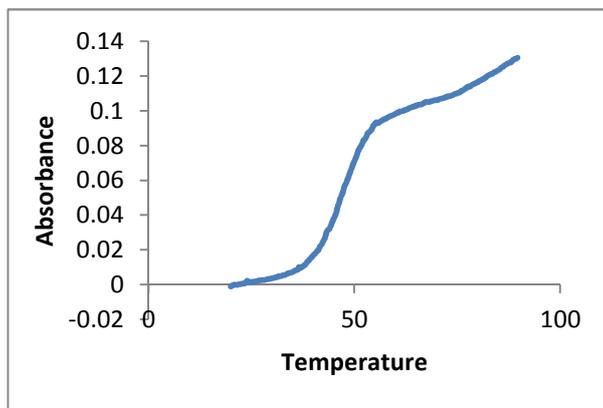
ODN1-1A
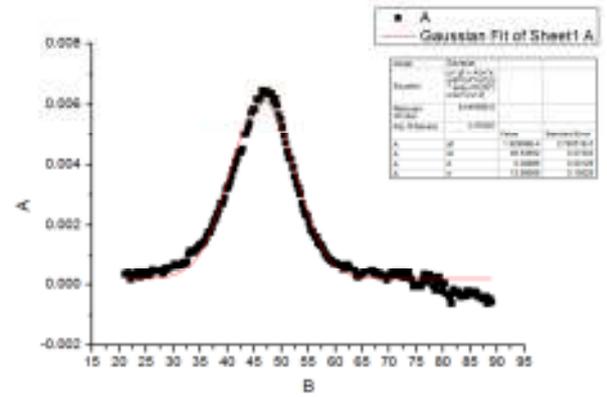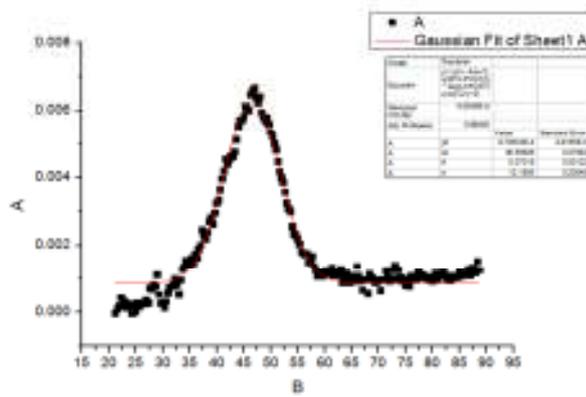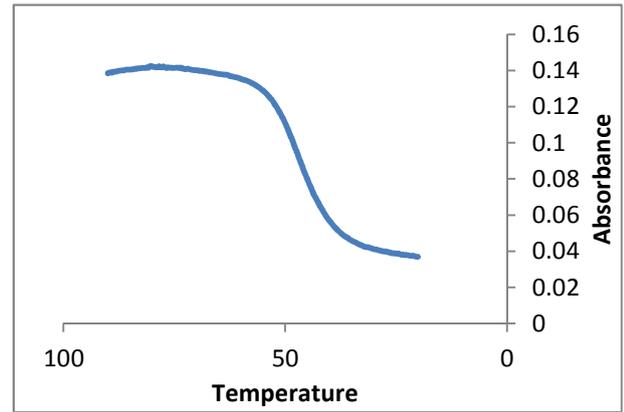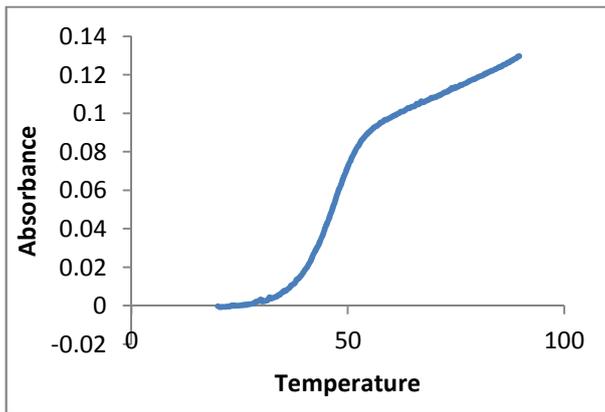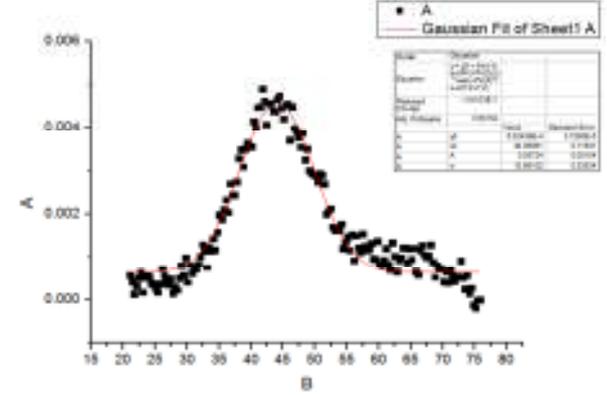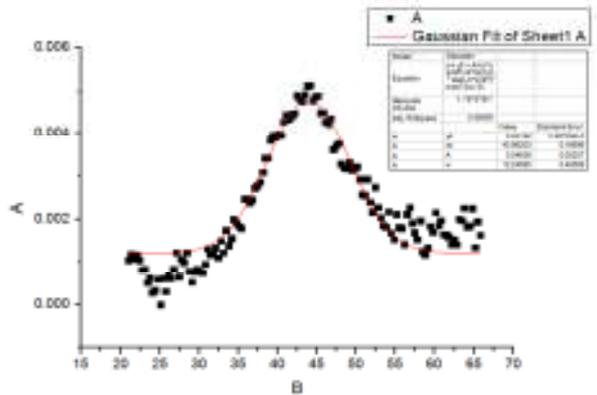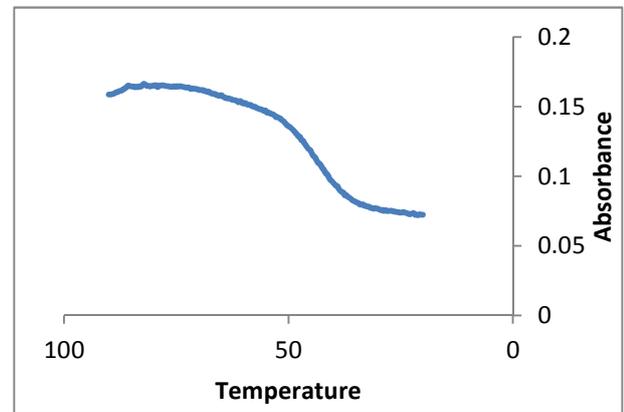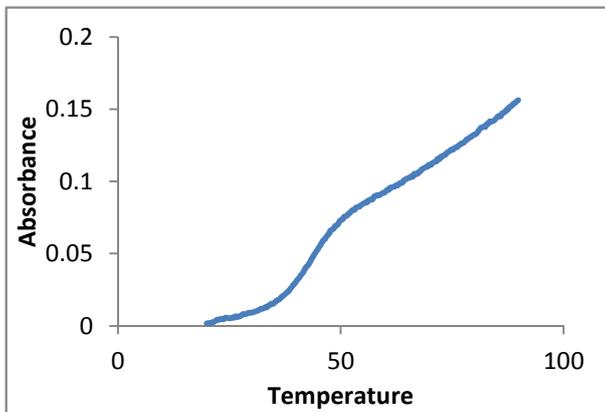
ODN1-1G

ODN2-2T
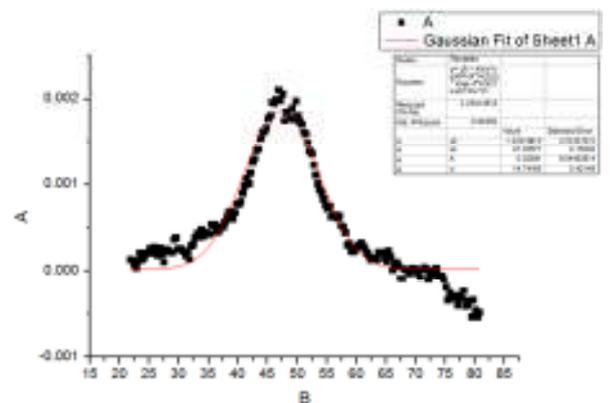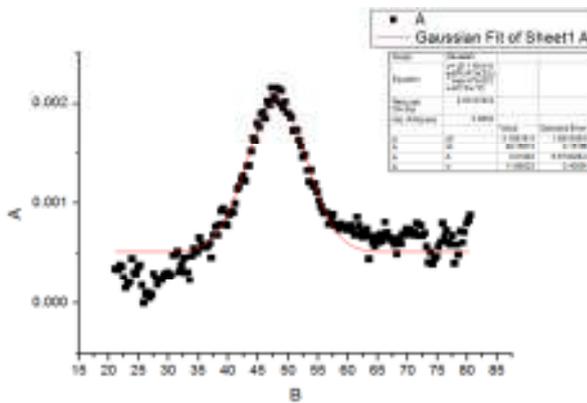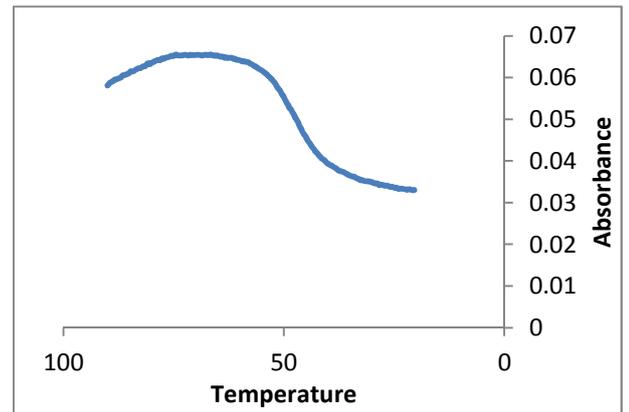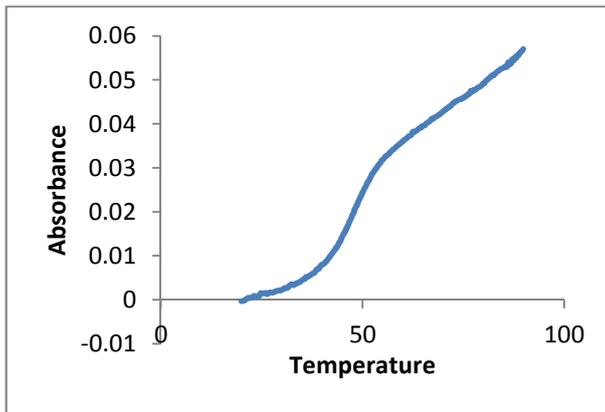
ODN1-2C
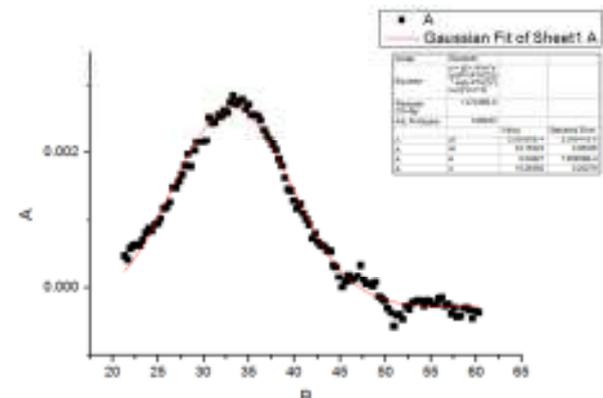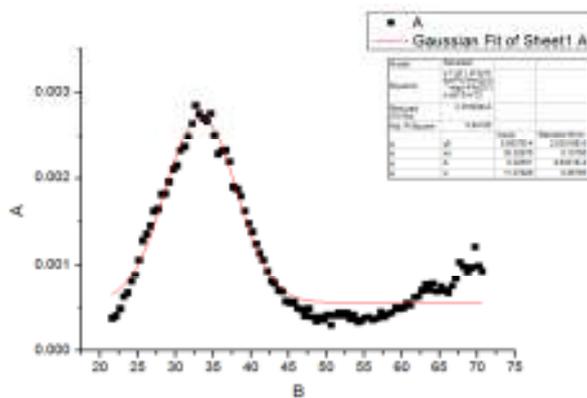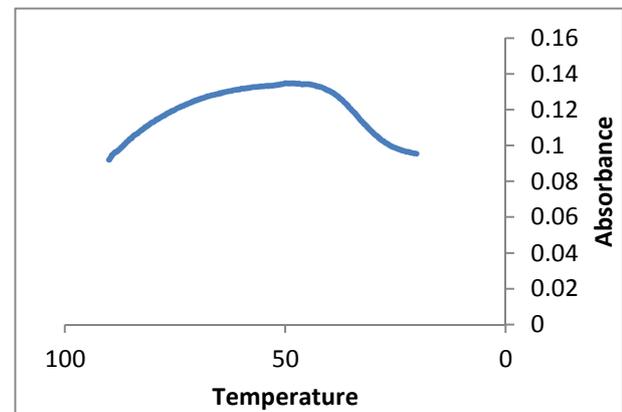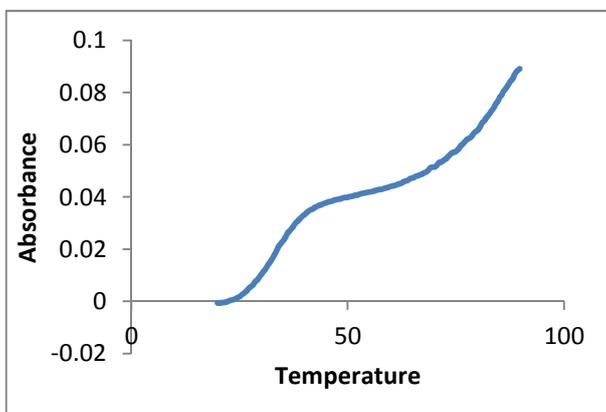
ODN2-2A
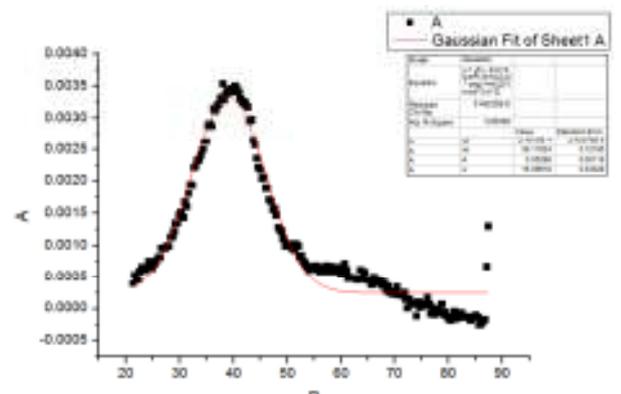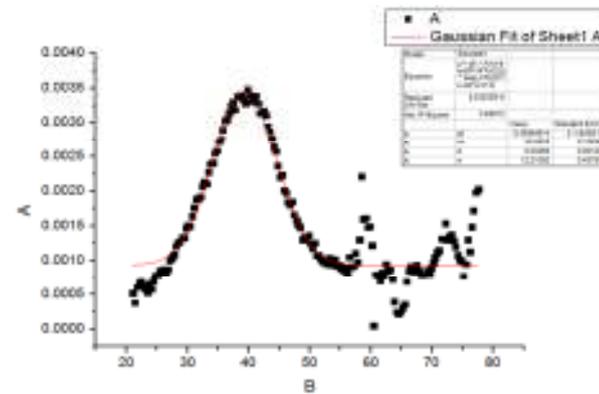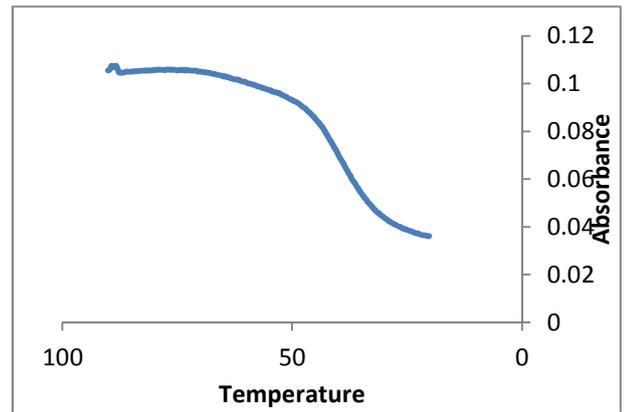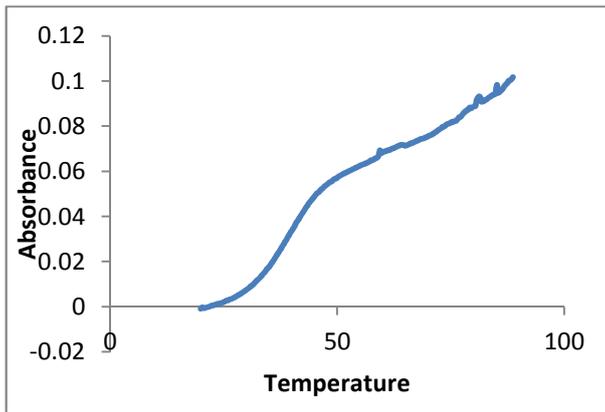
ODN2-2G
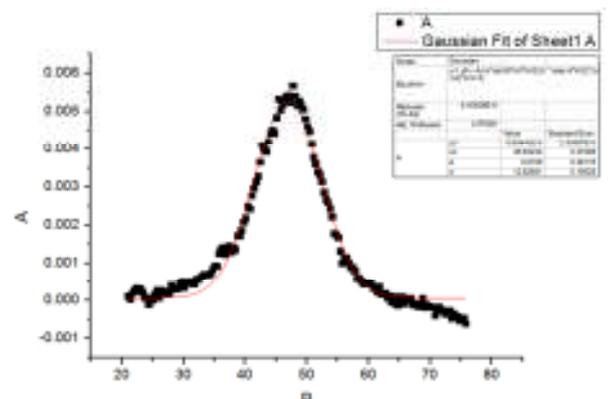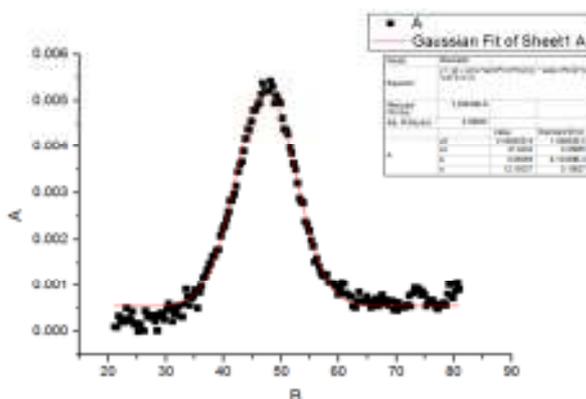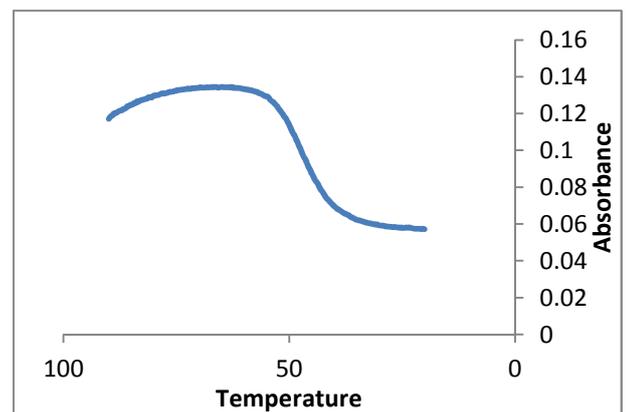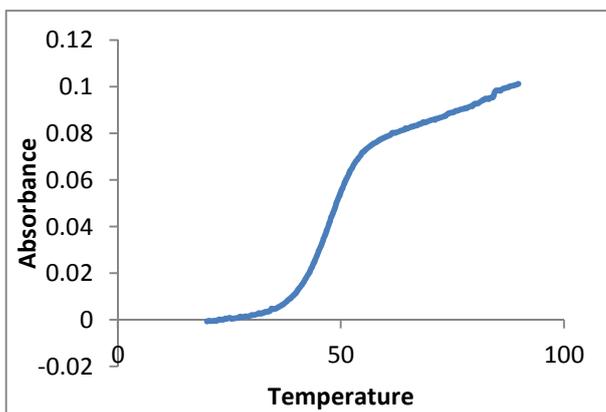
ODN3-3T
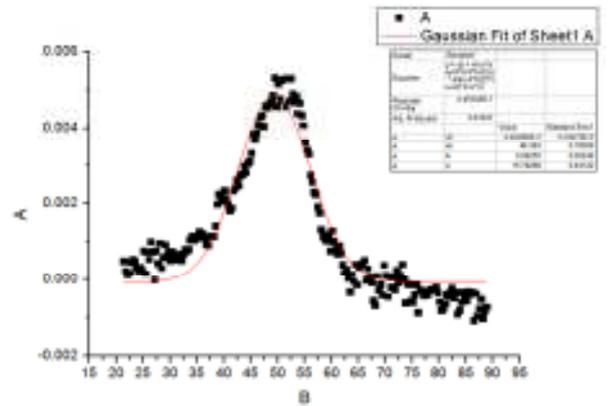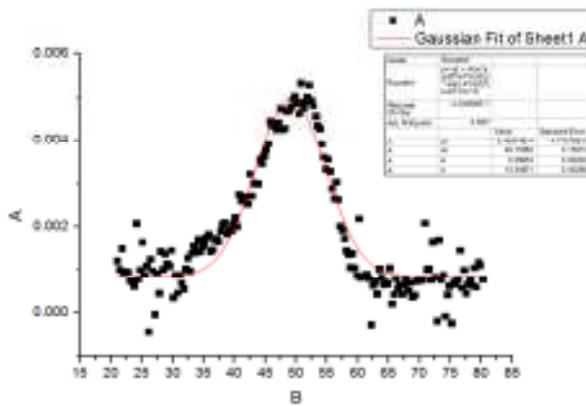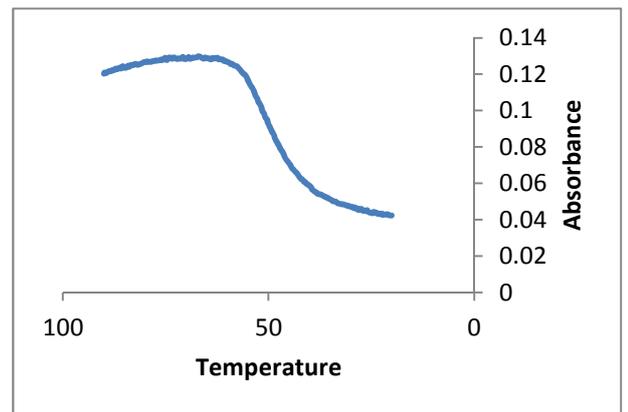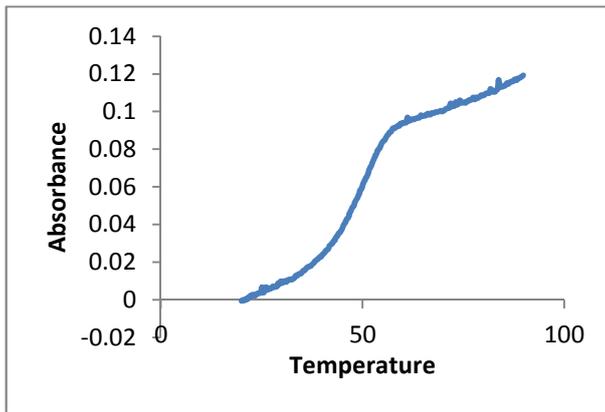
ODN3-3C

ODN3-3A

ODN3-3G

191
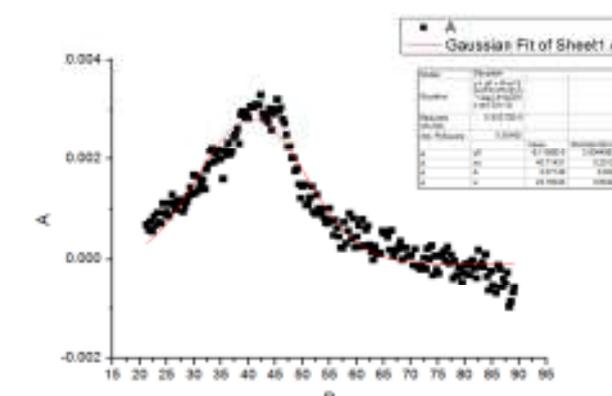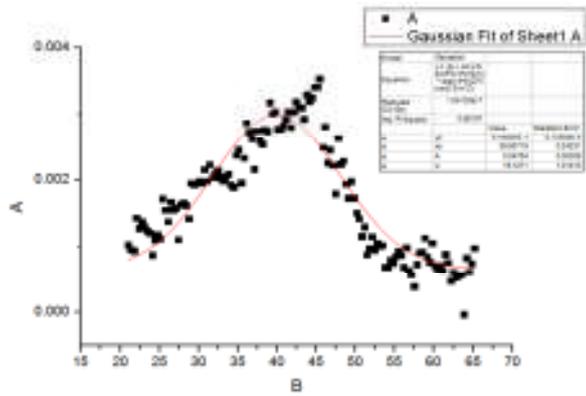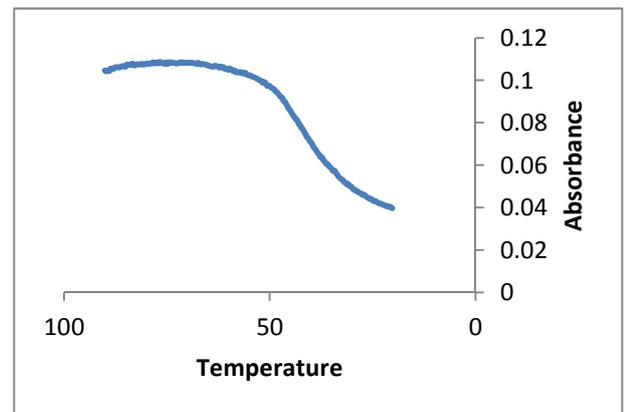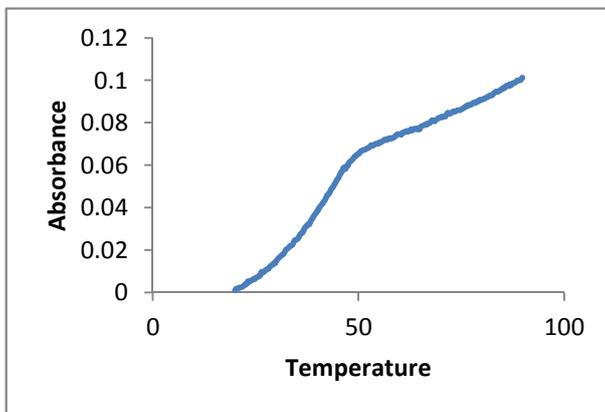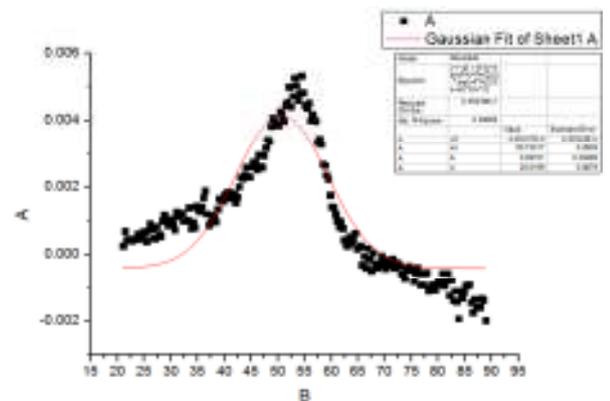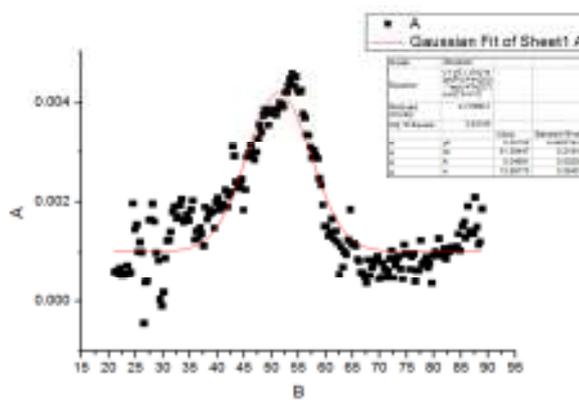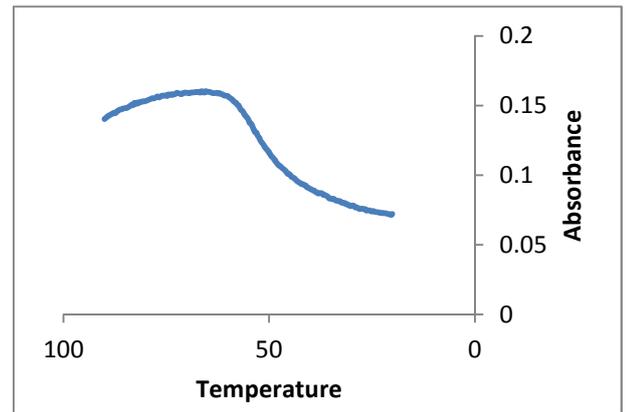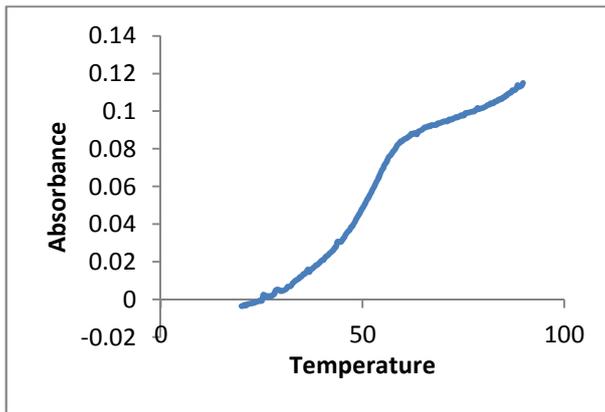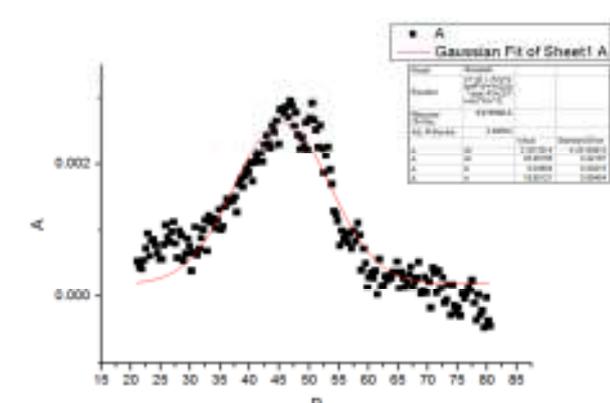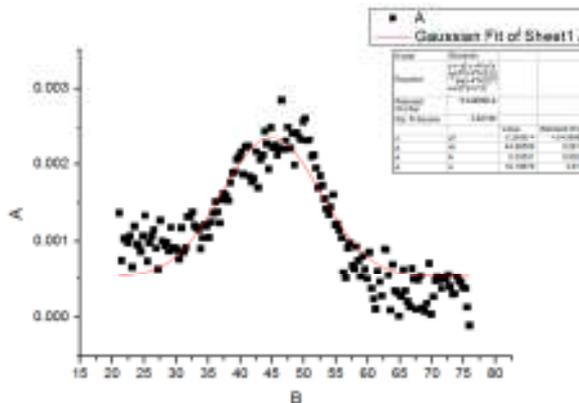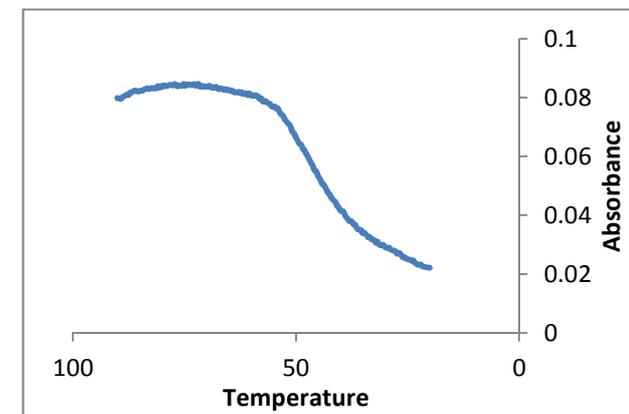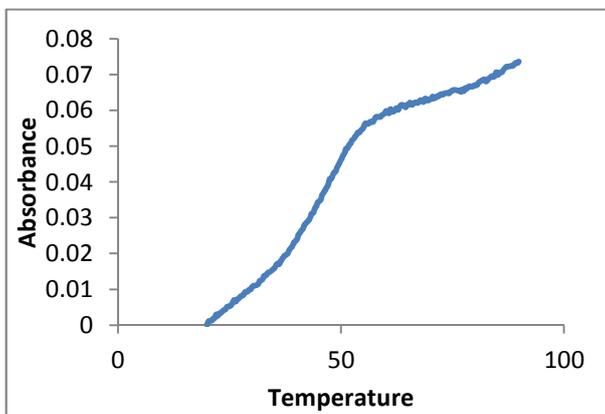
ODN4-4T

ODN4-4C

ODN4-4A

ODN4-4G

ODN5-5T

ODN5-5C

ODN5-5A

ODN5-5G
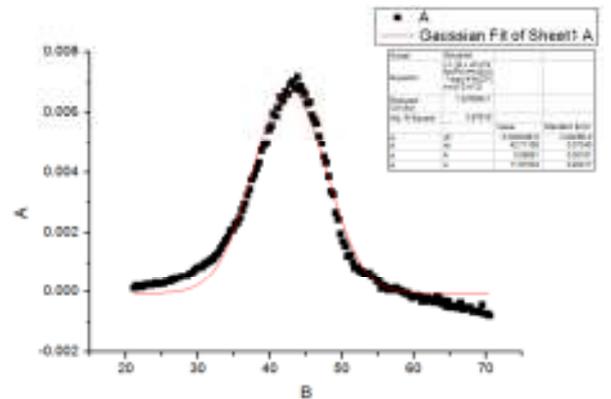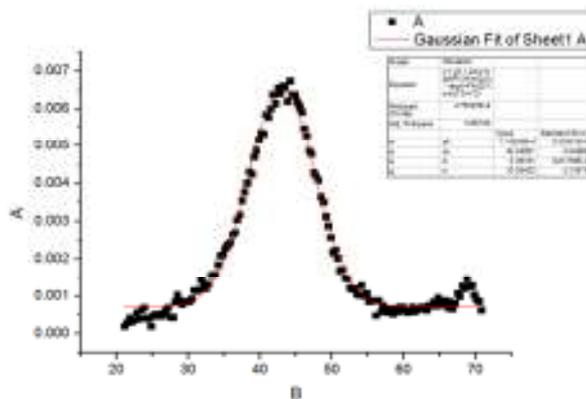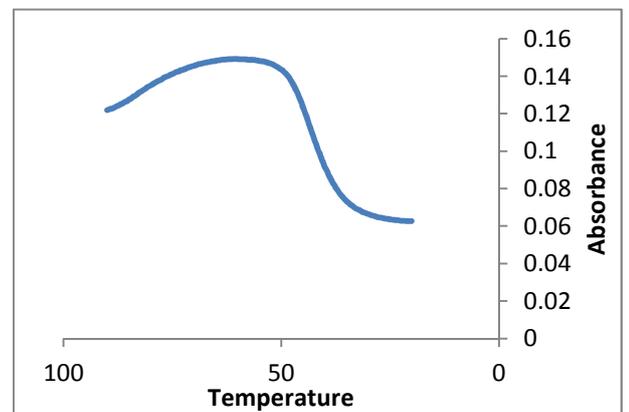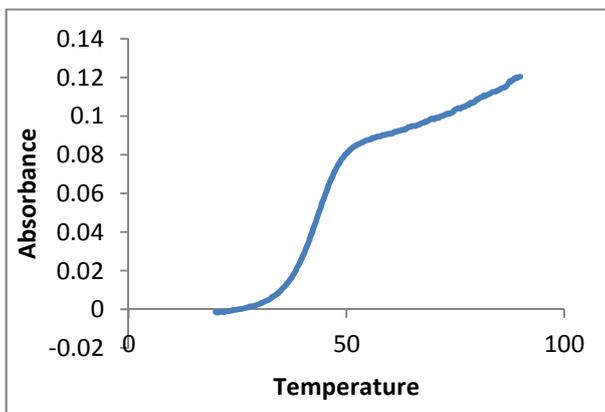
Appendix III - Circular Dichroism Spectra

ODN1-1T

ODN2-2T

Wavelength /nm

10 C —— 75 C

Wavelength /nm

C2-2T —— ODN2-2T

ODN3-3T

ODN4-4T

10 C   75 C

C4-4T   ODN4-4T

ODN5-5T

Top graph legend: 10 C — 75 C
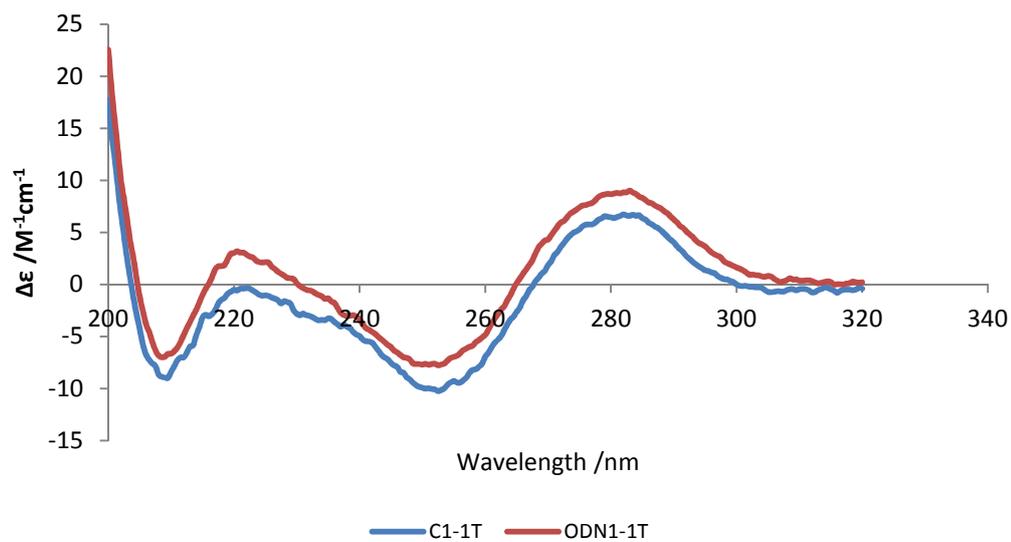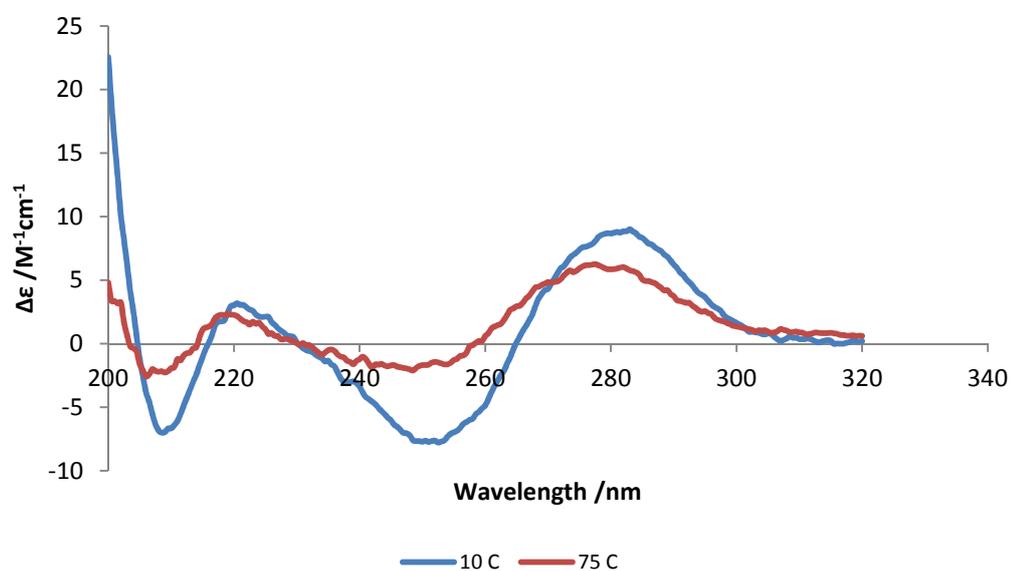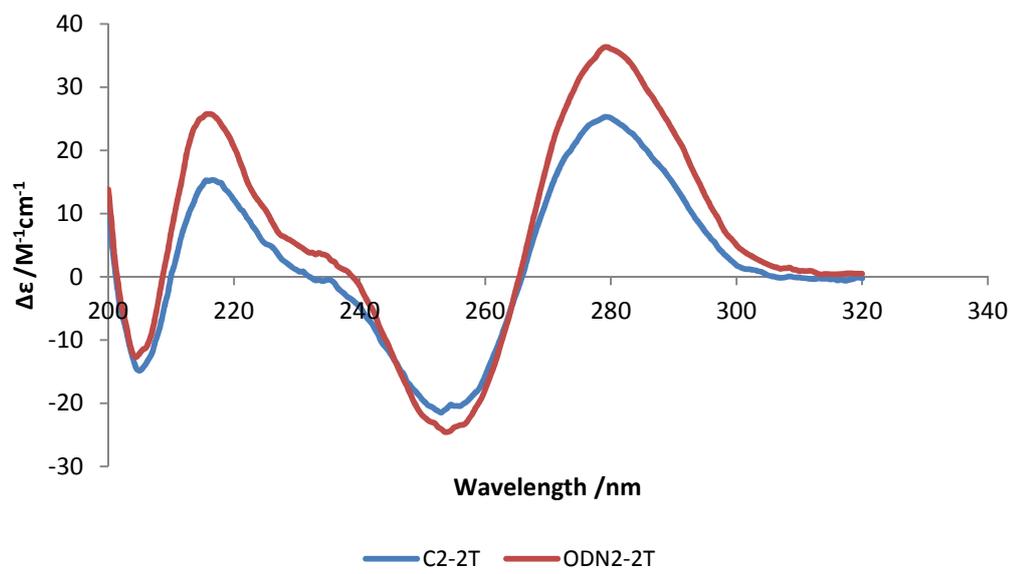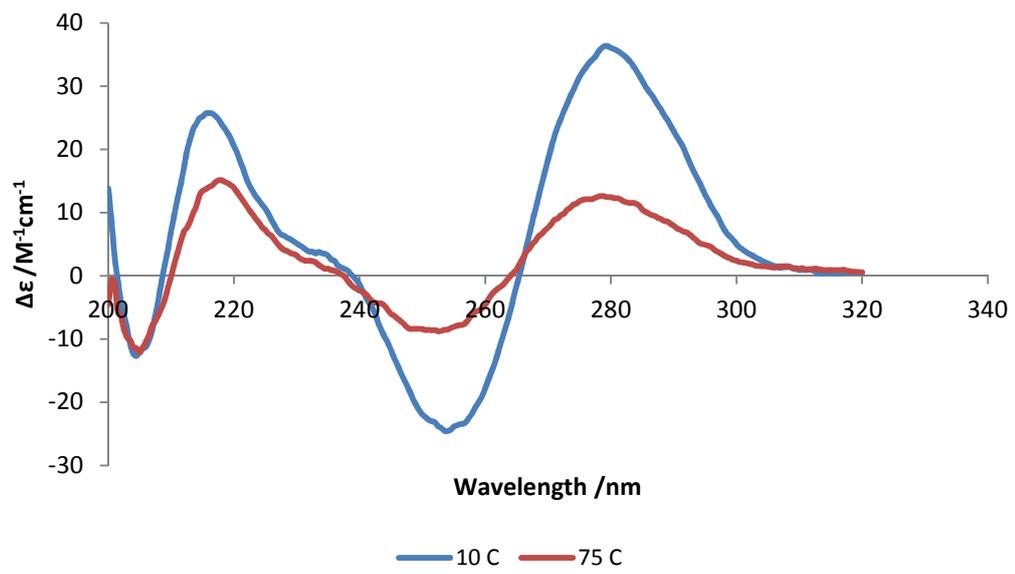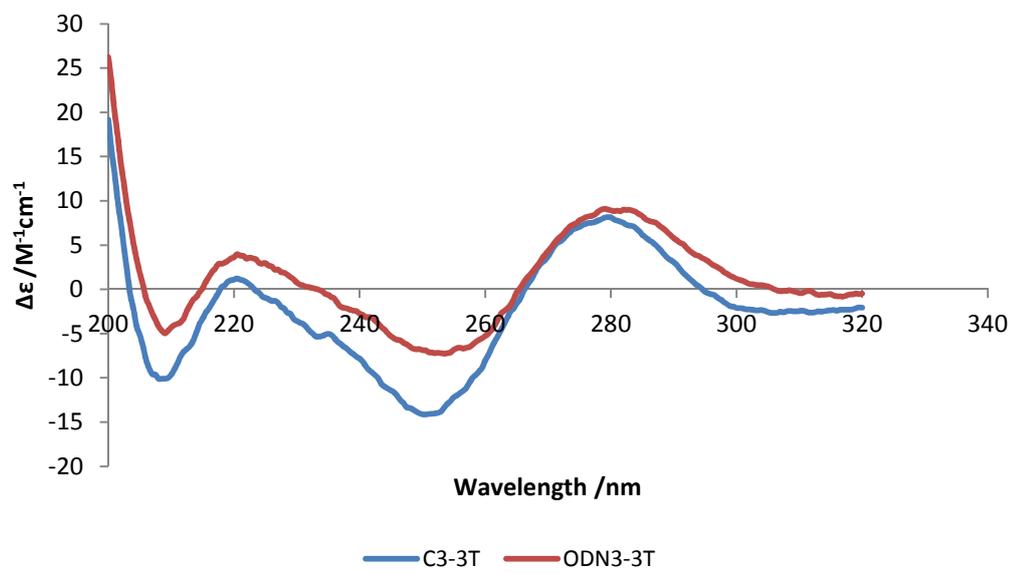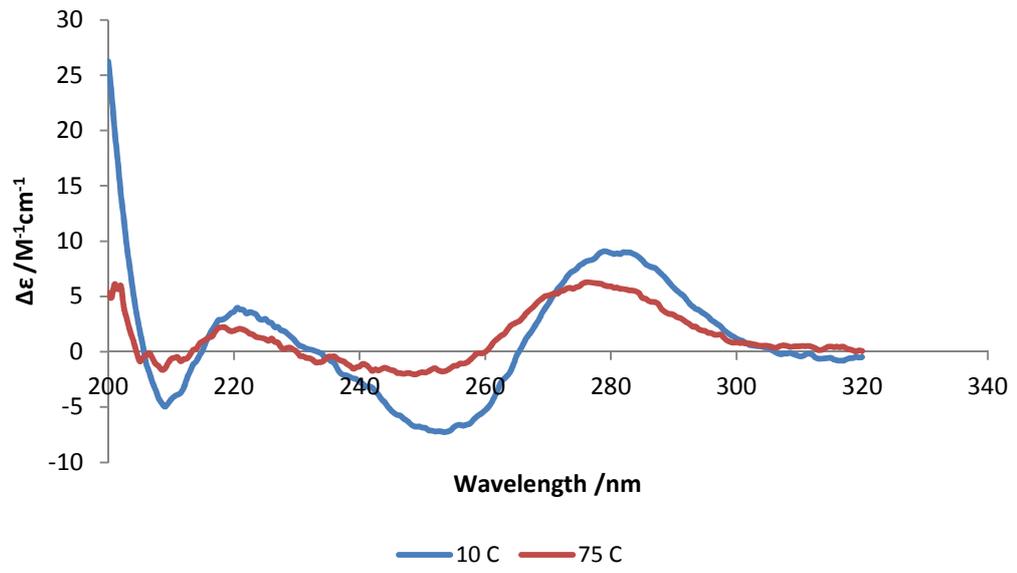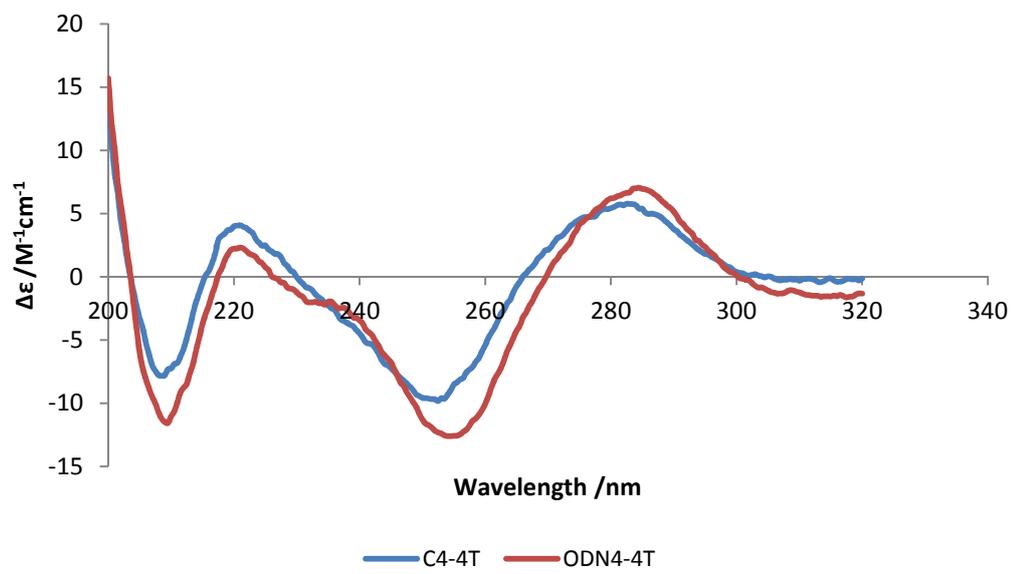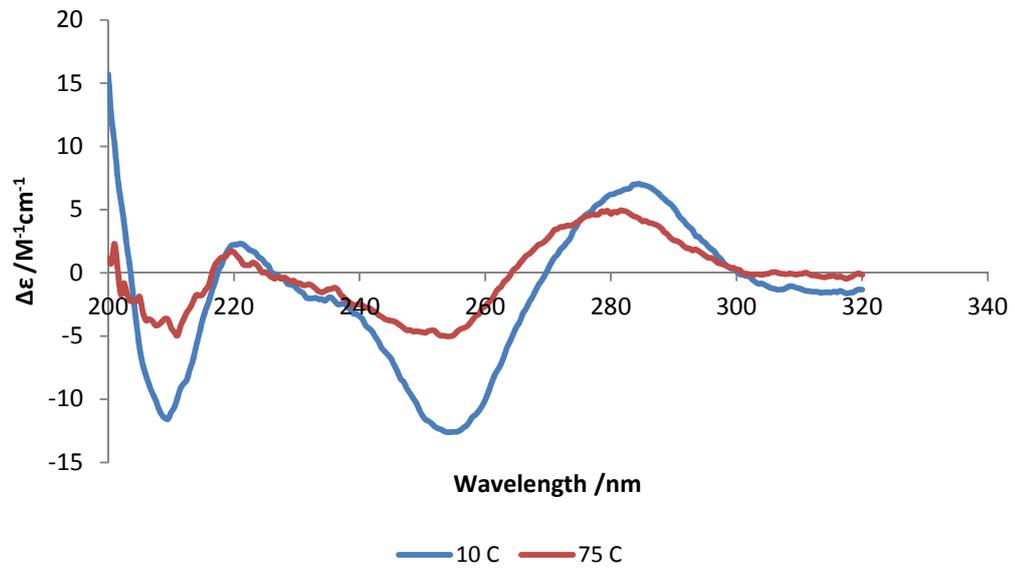
Bottom graph legend: C5-5T — ODN5-5T
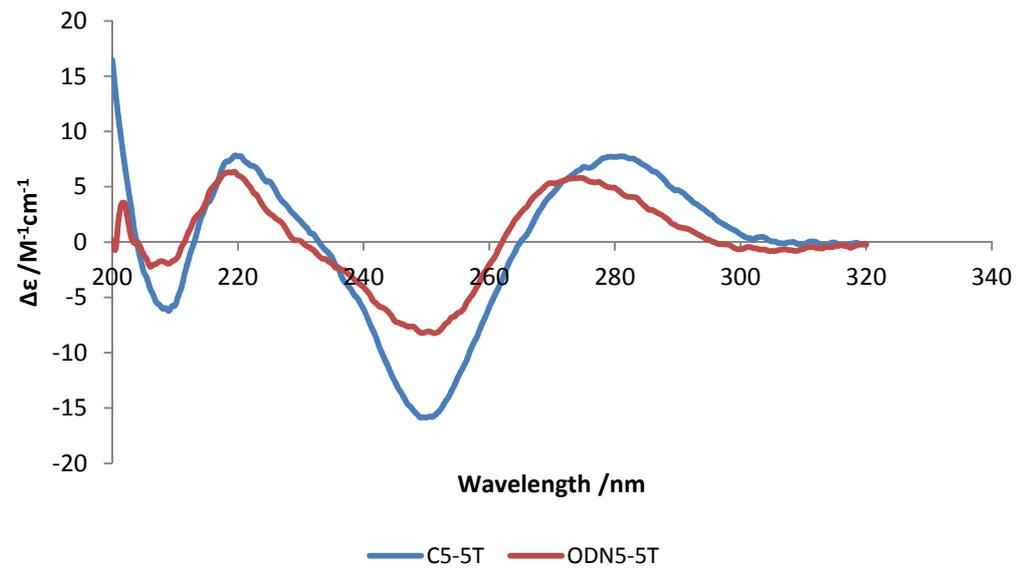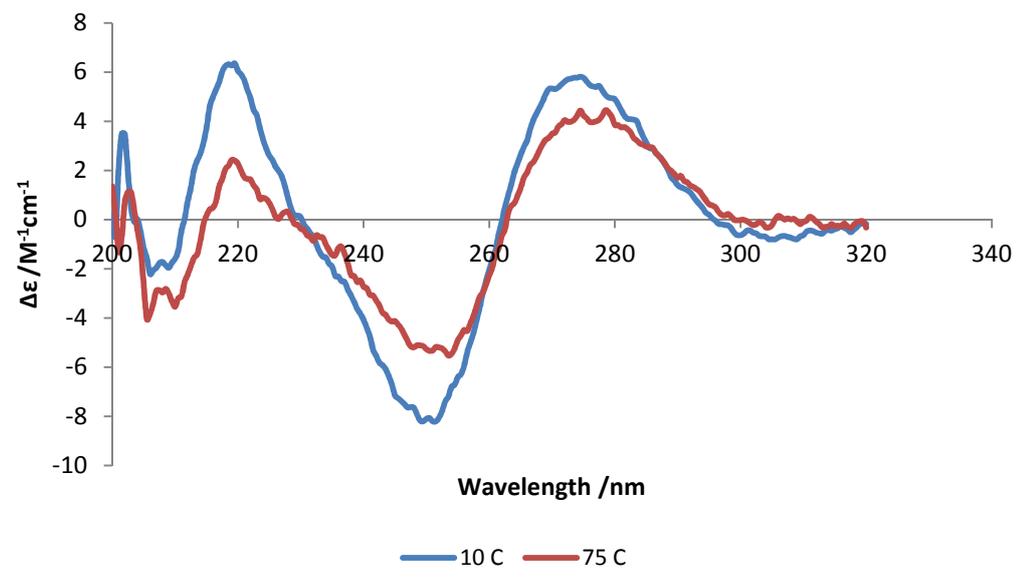
Appendix IV – Fluorescence spectroscopy of the oligonucleotides

Stern-Volmer plots

Single-stranded oligonucleotides
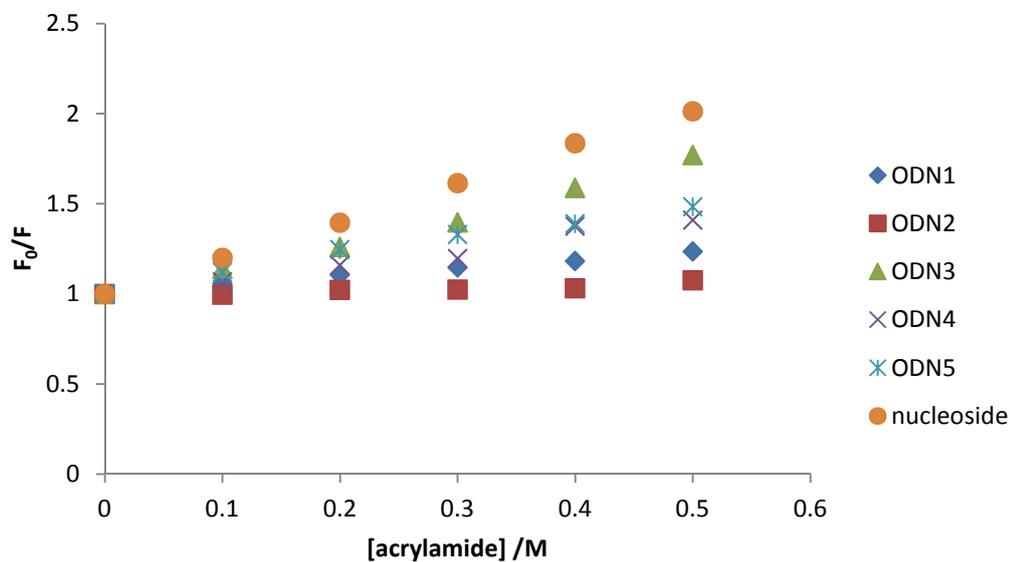
Double-stranded oligonucleotides
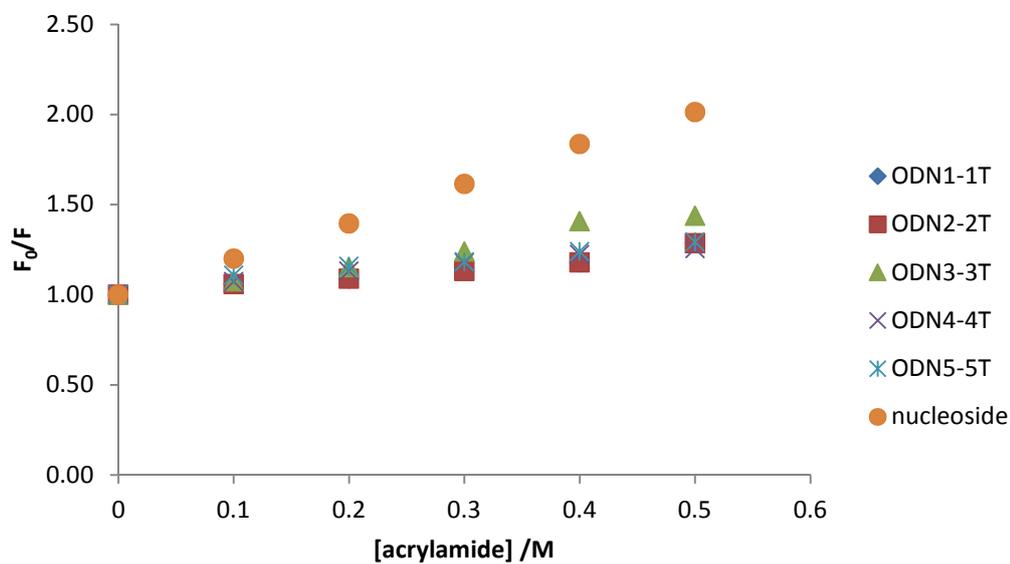
Fluorescence titrations (representative spectra)

ODN1-1T

ODN1-1C

ODN1-1A

ODN1-1G

◆ Fluorescence emission   ■ Emission maximum
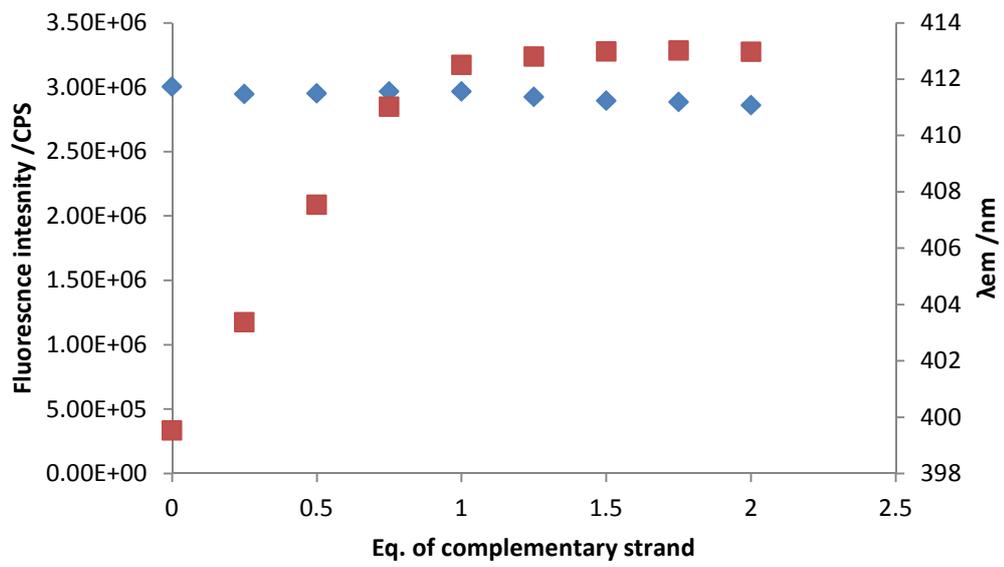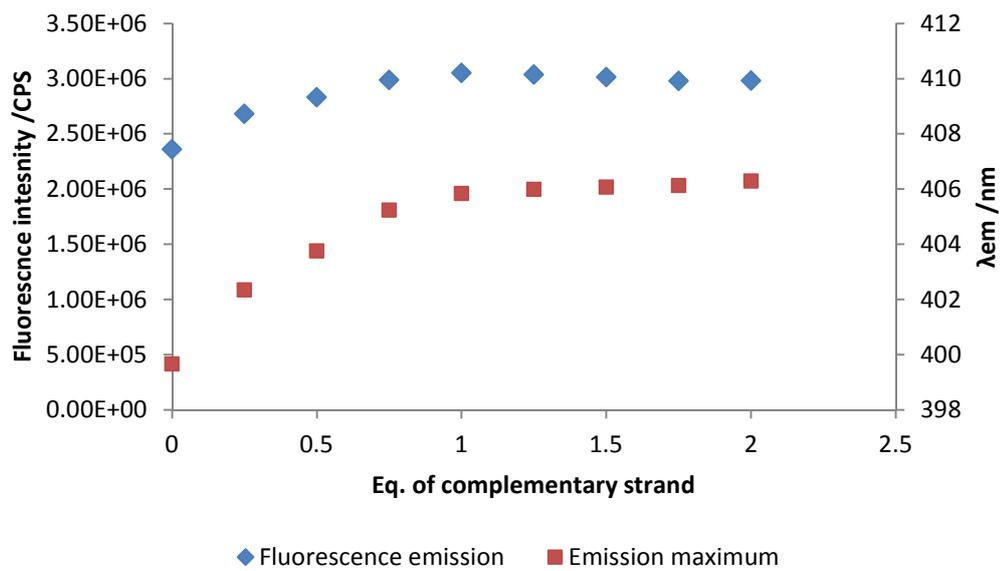
ODN2-2T

ODN2-2C

ODN2-2A

ODN2-2G

ODN3-3T

Fluorescence emission ◆   **Emssion maximum** ■

ODN3-3C

Fluorescence emission ◆   **Emssion Maximum** ■

ODN3-3A

◆ Fluorescence emission　■ Emssion Maximum

ODN3-3G

ODN4-4T

ODN4-4C

ODN4-4A

ODN4-4G

ODN5-5T

ODN5-5C

ODN5-5A

Fluorescence emission ◆   **Emission maximum** ■

ODN5-5G

Fluorescence emission ◆   **Emission maximum** ■
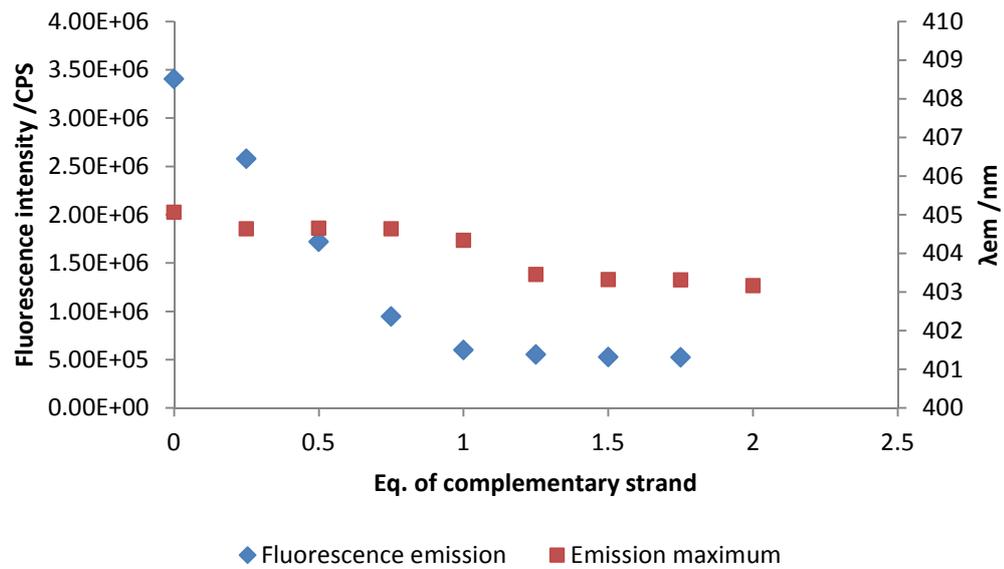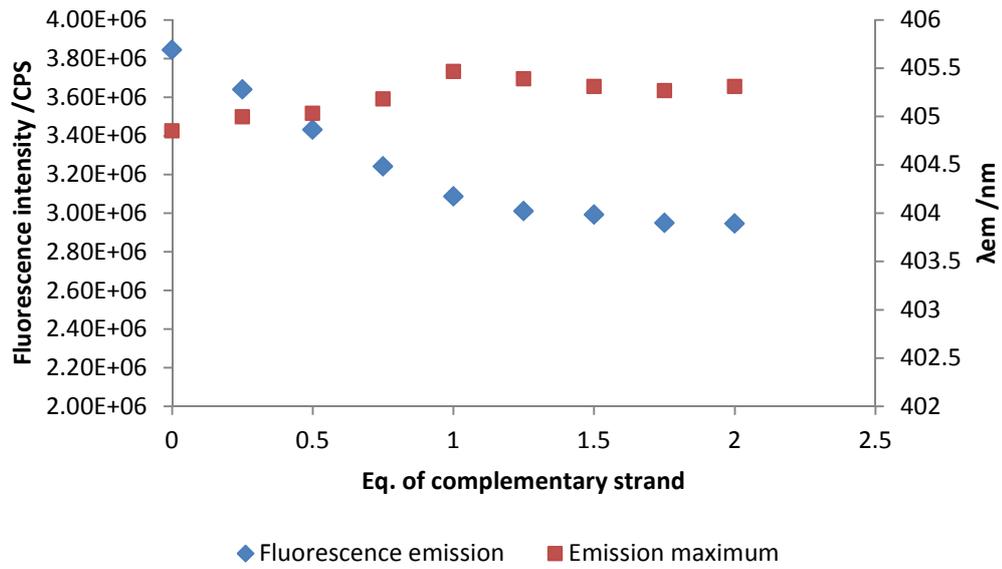
Appendix V - pH dependence of oligonucleotides
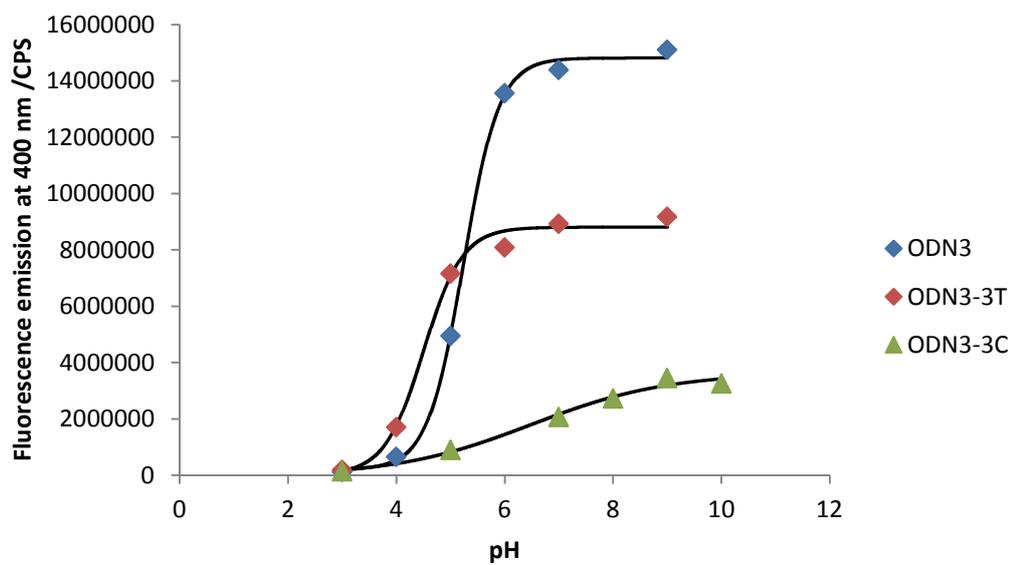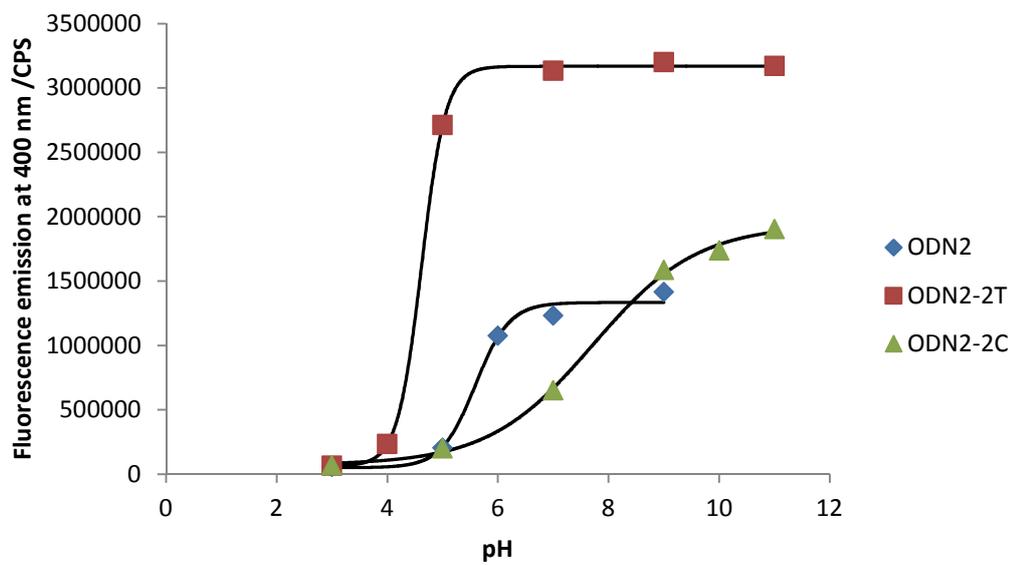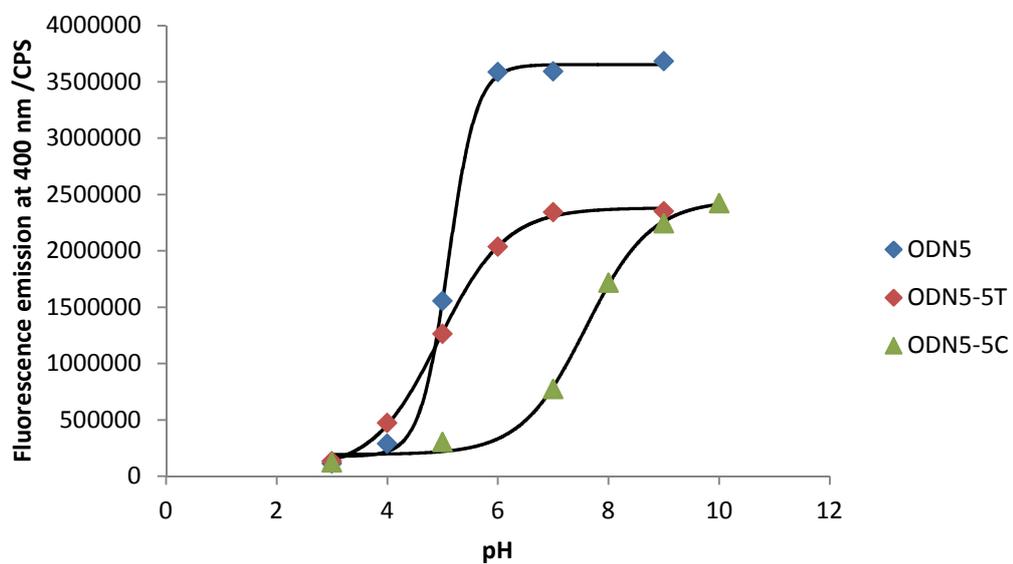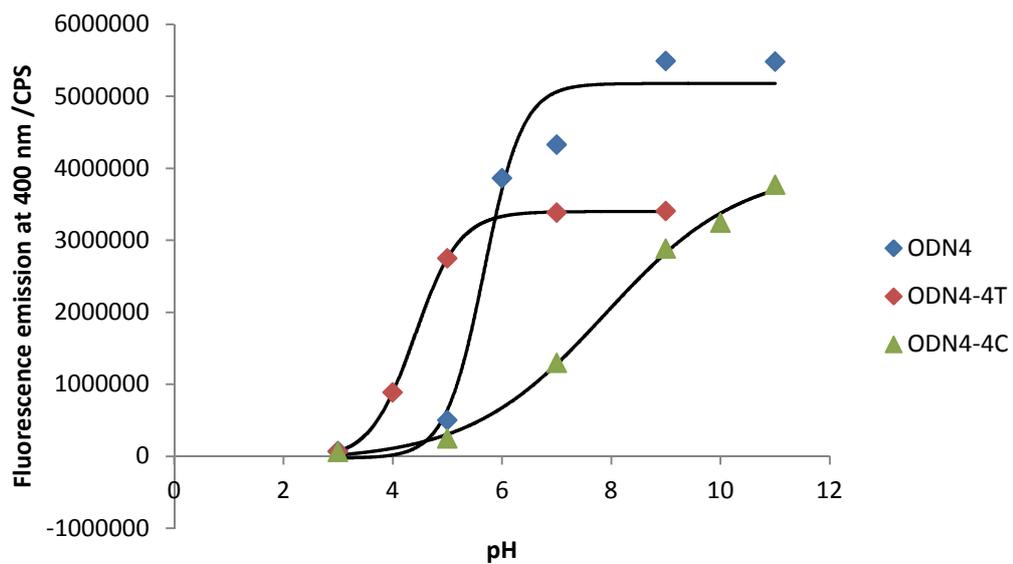
Appendix VI - NMR Spectra

13 - ^{1}H NMR DMSO-d_6
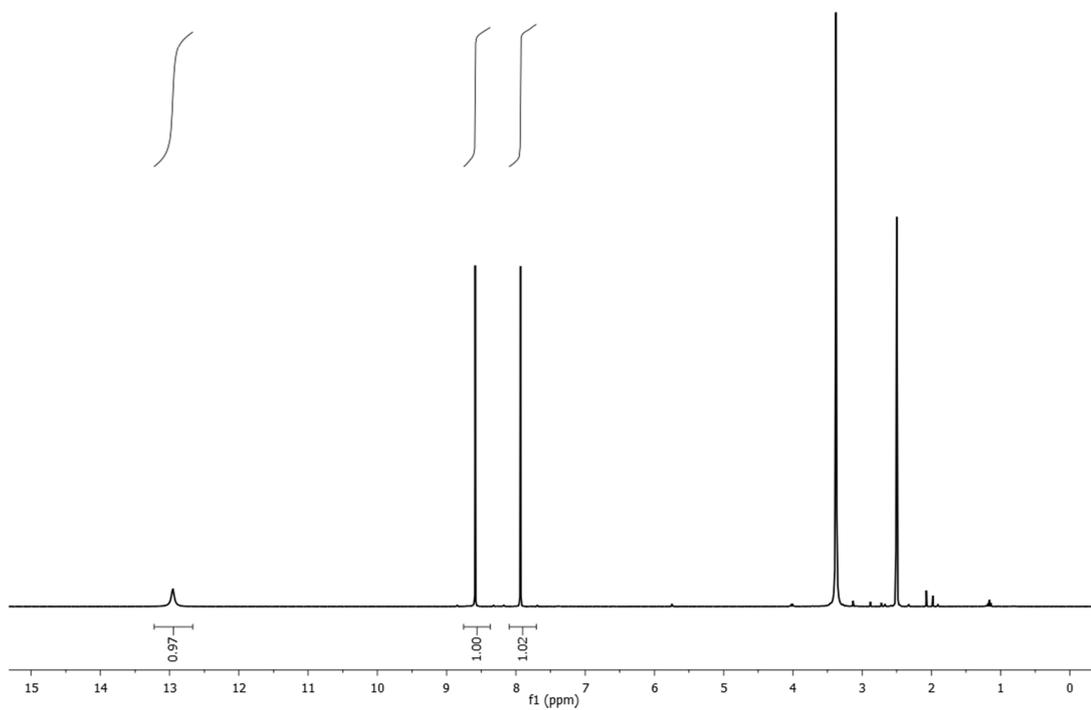
13 – ^{13}C NMR DMSO-d_6

21 - ¹H NMR acetone-*d*₆

7.41
7.39
7.24
7.22

1.01
1.04

10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
f1 (ppm)

21 – ¹³C NMR acetone-*d*₆

134.67
134.66
134.64
134.62
129.91
119.68
119.67
119.66

220 210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -10 -20
f1 (ppm)

215

21 – ^{19}F NMR acetone-d_6

142.90
143.01
143.15
143.29
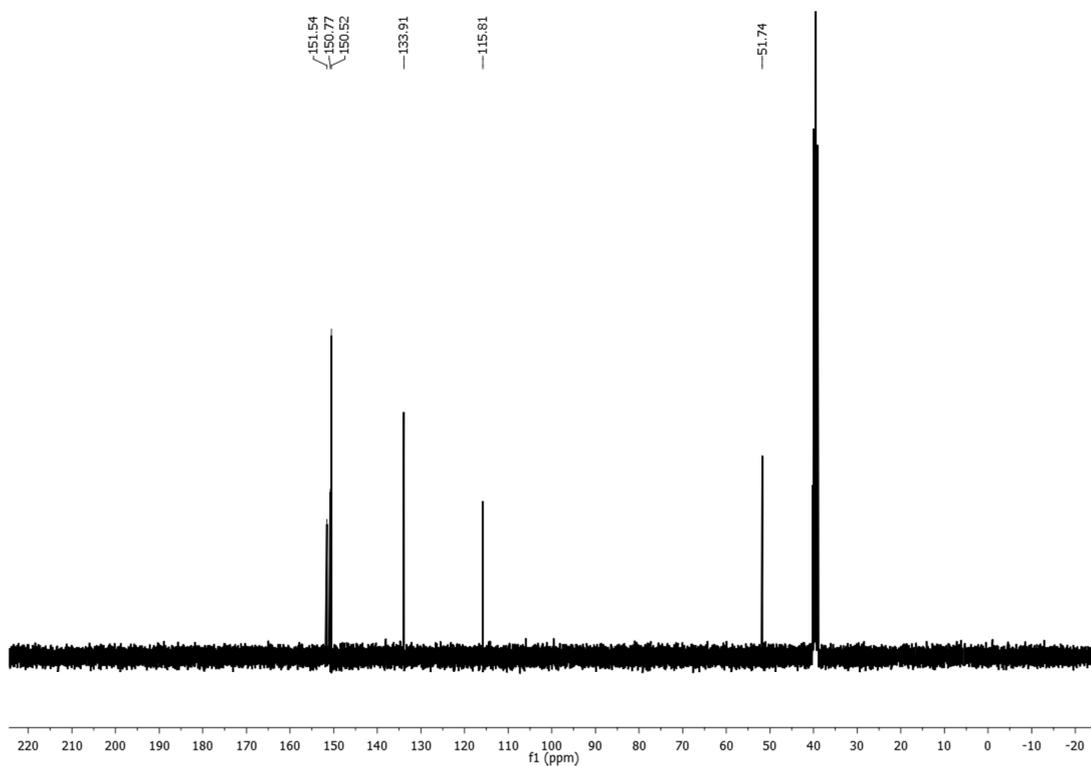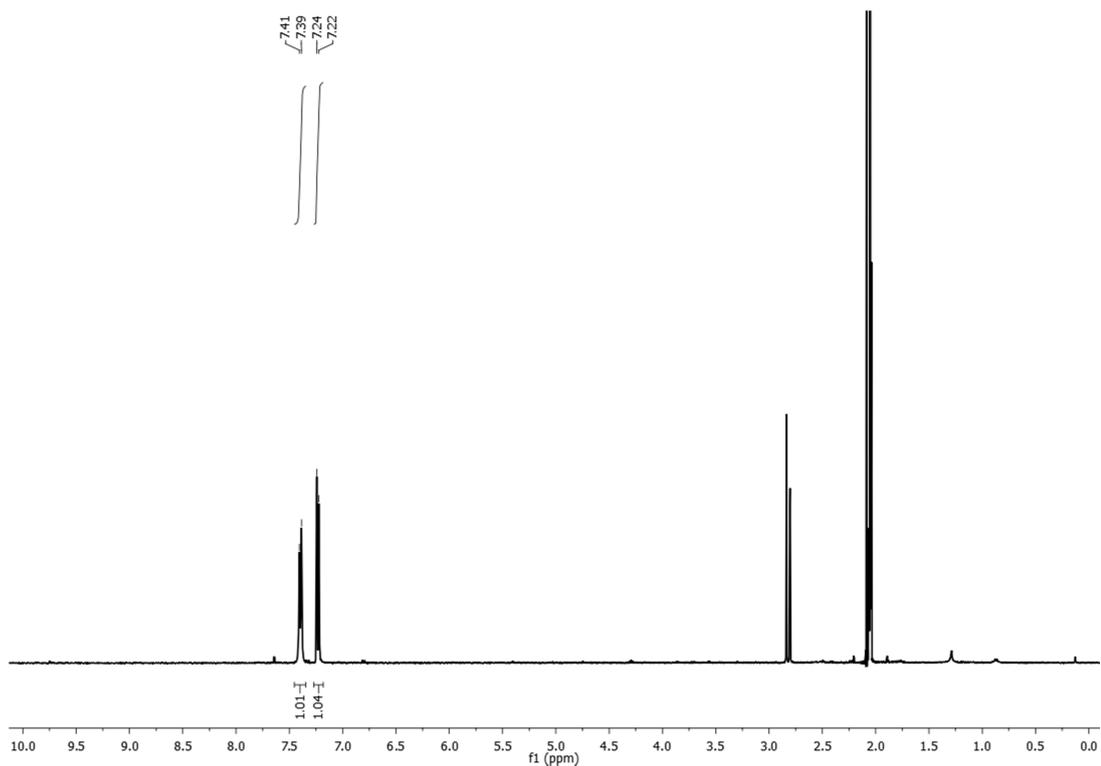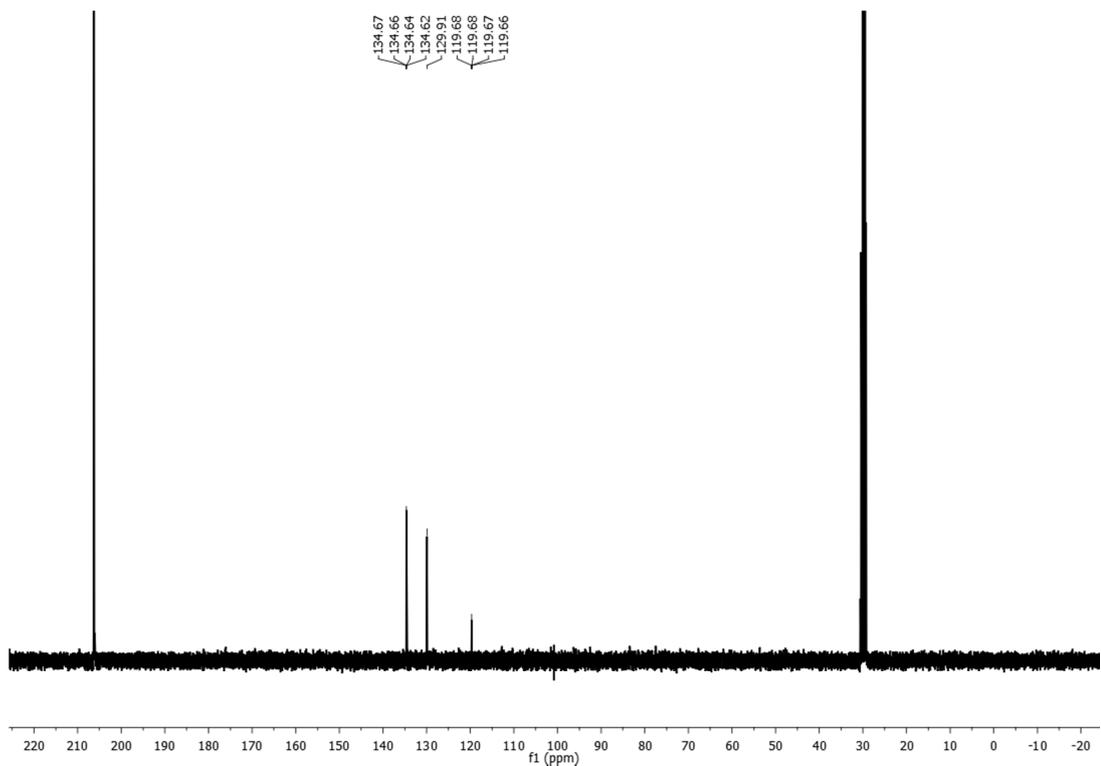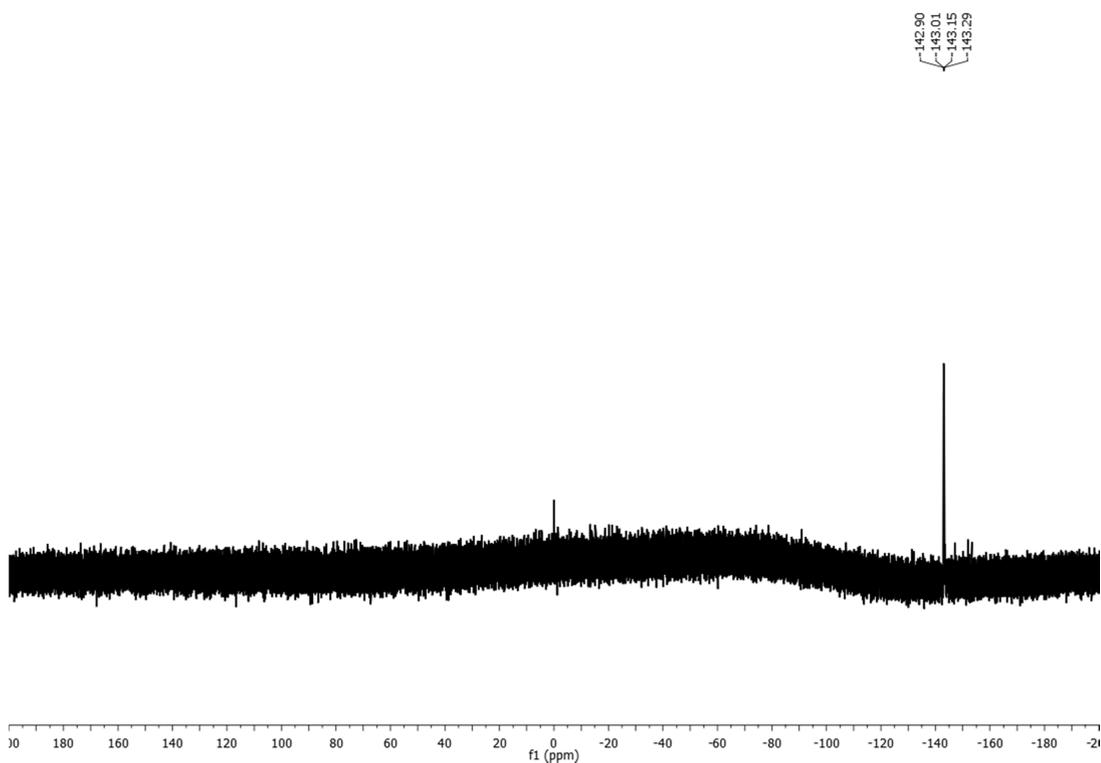
21 – ^{11}B NMR acetone-d_6

2.84
2.40
2.01
1.63
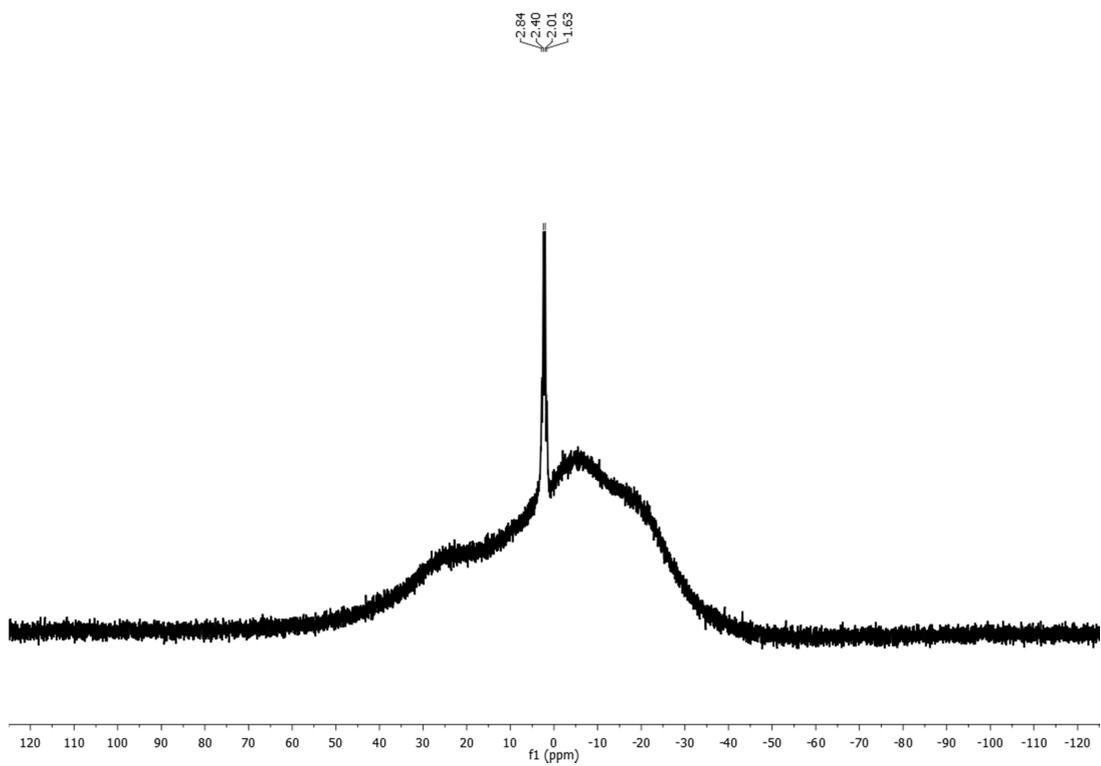
26 – ¹H NMR CDCl₃

26 – ¹³C NMR CDCl₃

217

26 – ¹¹B NMR CDCl₃

—25.64

120 110 100 90 80 70 60 50 40 30 20 10 0 -10 -20 -30 -40 -50 -60 -70 -80 -90 -100 -110 -120
f1 (ppm)

28 – ¹H NMR CDCl₃

7.73
7.71
7.45
7.43

—3.76

—1.02

—0.25

2.04 2.00   4.09   6.04   9.07

8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
f1 (ppm)

28 – ^{13}C NMR CDCl$_3$

133.68
131.17
125.23
105.55
100.06
95.22
72.47
32.03
22.06
0.12

190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
f1 (ppm)

28 – ^{11}B NMR CDCl$_3$

26.43

120 110 100 90 80 70 60 50 40 30 20 10 0 -10 -20 -30 -40 -50 -60 -70 -80 -90 -100 -110 -120
f1 (ppm)

27a – ¹H NMR CDCl₃

7.79
7.77
7.55
7.53
7.52
7.51
7.35
7.34
7.33

3.78

1.03

2.10
4.19
3.12

4.44

6.00

10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
f1 (ppm)

27a – ¹³C NMR CDCl₃

133.84
131.78
130.83
128.48
128.42
125.44
123.42

90.46
89.89

72.49

32.05
22.06

220 210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -10 -20
f1 (ppm)

220

27a – ^{11}B NMR CDCl$_3$

29 – ^{1}H NMR CDCl$_3$
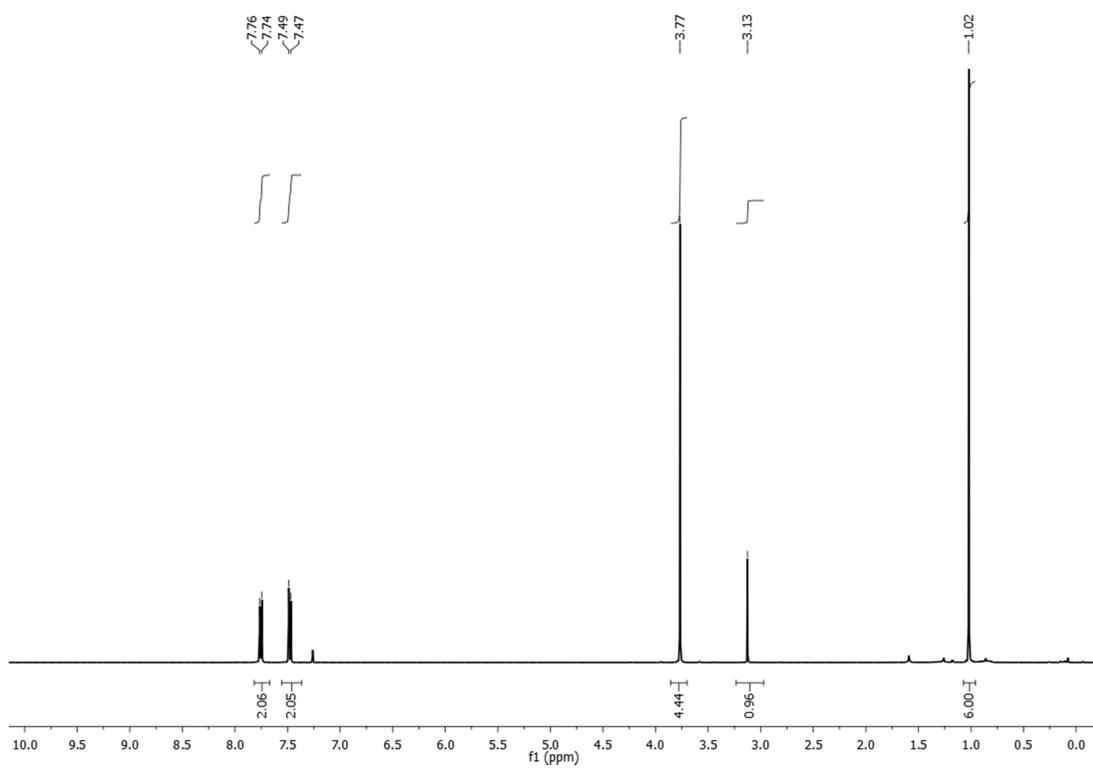
29 – ^{13}C NMR CDCl$_3$
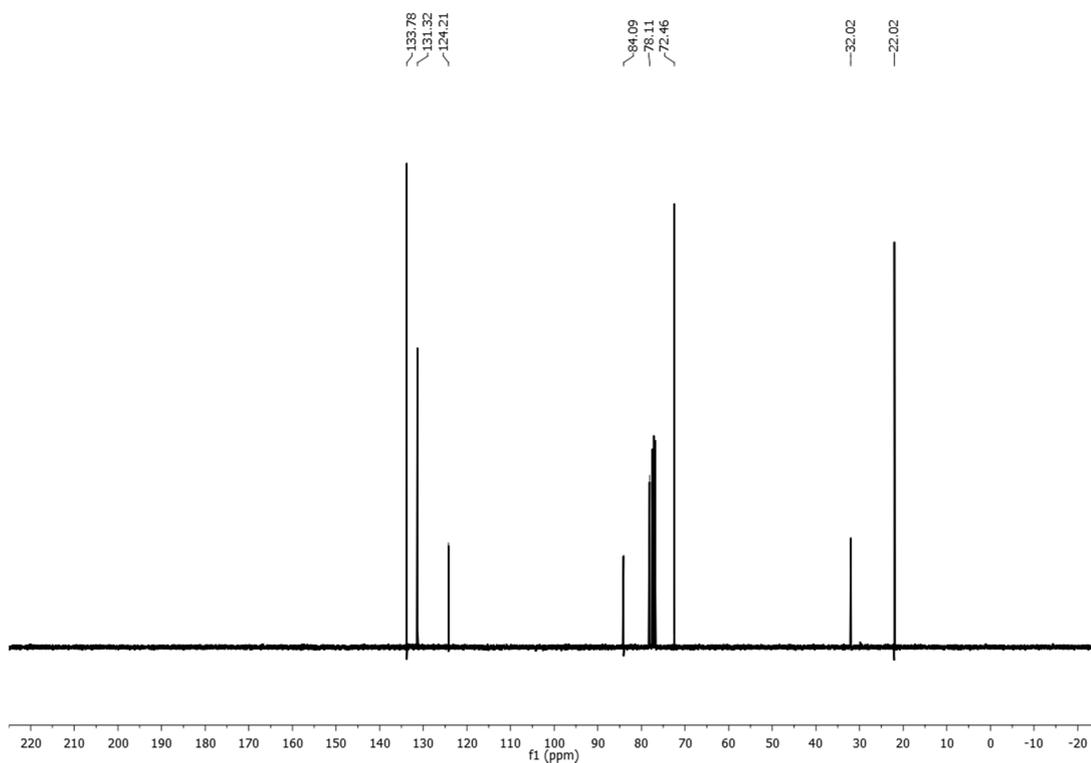
29 – ^{11}B NMR CDCl$_3$

27d – ¹H NMR CDCl₃

7.77
7.75
7.50
7.48
7.46
6.89
6.87
3.83
3.77
1.03

2.08
4.06
2.07
3.01
3.88
6.00

10.0   9.5   9.0   8.5   8.0   7.5   7.0   6.5   6.0   5.5   5.0   4.5   4.0   3.5   3.0   2.5   2.0   1.5   1.0   0.5   0.0
f1 (ppm)

27d – ¹³C NMR CDCl₃

159.79
133.82
133.24
130.66
125.81
115.56
114.14
90.52
88.63
72.49
55.45
32.05
22.06

220   210   200   190   180   170   160   150   140   130   120   110   100   90   80   70   60   50   40   30   20   10   0   -10   -20
f1 (ppm)

27d – ^{11}B NMR CDCl$_3$

−26.66

120 110 100 90 80 70 60 50 40 30 20 10 0 −10 −20 −30 −40 −50 −60 −70 −80 −90 −100 −110 −120
f1 (ppm)

27d – 1H NMR CDCl$_3$

7.78
7.76
7.53
7.52
7.52
7.51
7.49
7.31
7.30
7.30
7.29
7.21
7.21
7.20
7.19

−3.77

−1.03

2.00
0.98
1.99
1.01
1.00

4.36

6.04

10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
f1 (ppm)

27d – ^{13}C NMR CDCl$_3$

27d – ^{11}B NMR CDCl$_3$
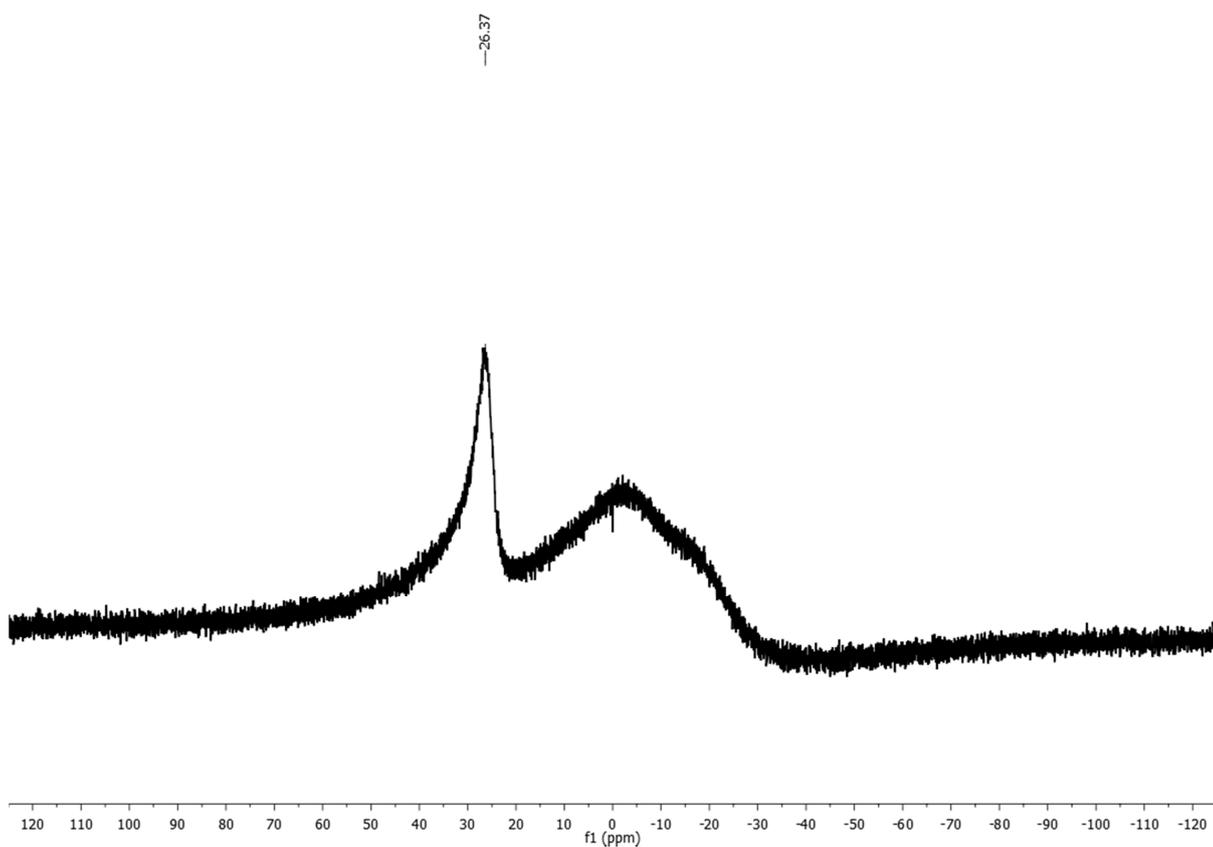
27e– ^{1}H NMR CDCl$_3$
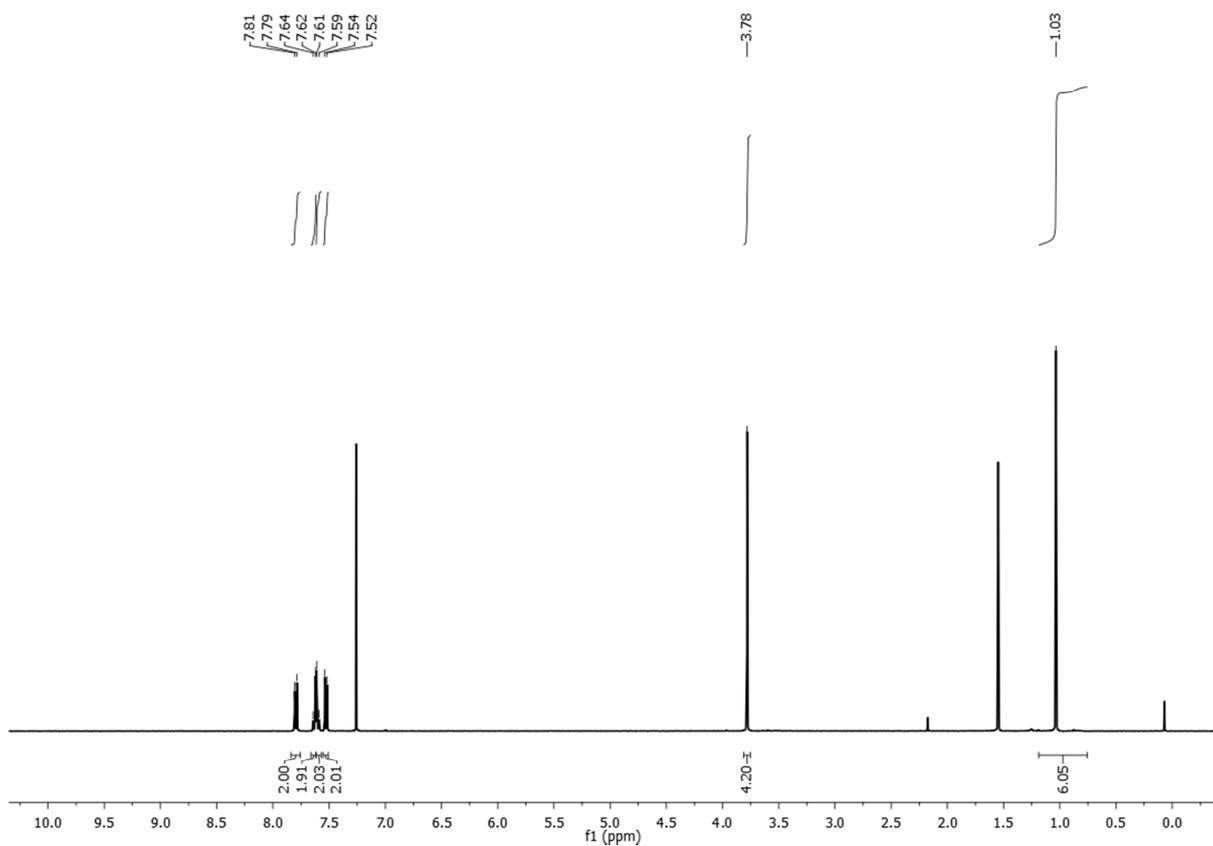
27e – ^{13}C NMR CDCl$_3$

27e– ^{11}B NMR CDCl$_3$

—25.91

120 110 100 90 80 70 60 50 40 30 20 10 0 -10 -20 -30 -40 -50 -60 -70 -80 -90 -100 -110 -120
f1 (ppm)
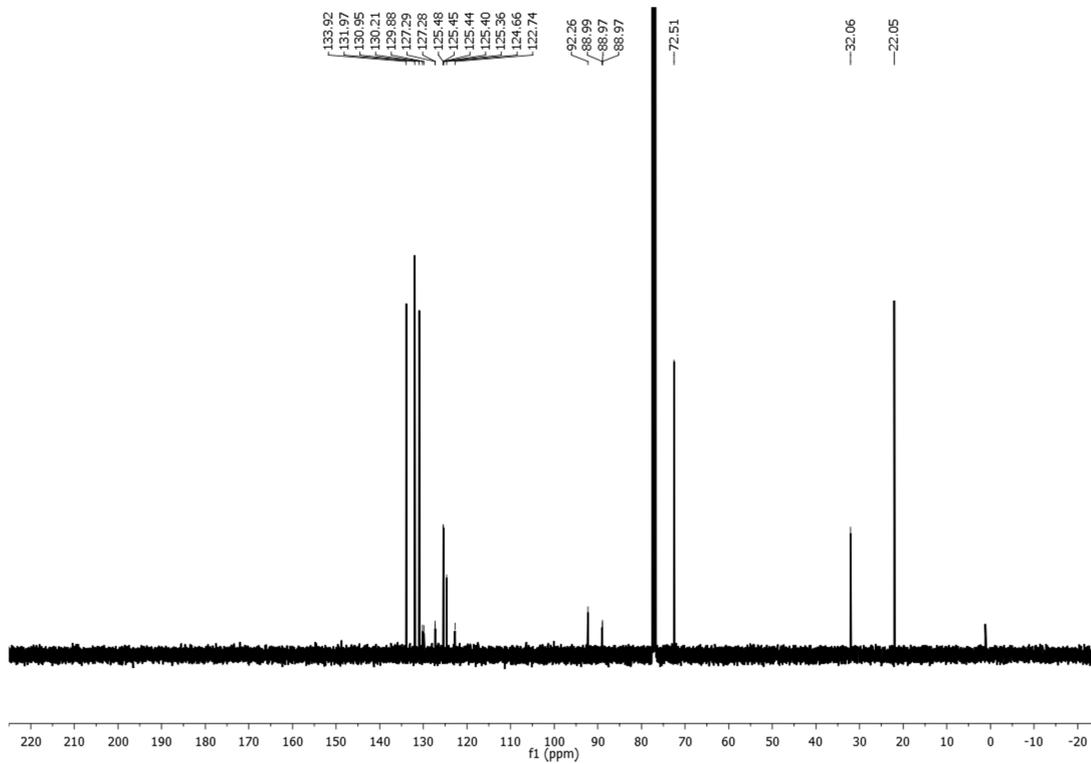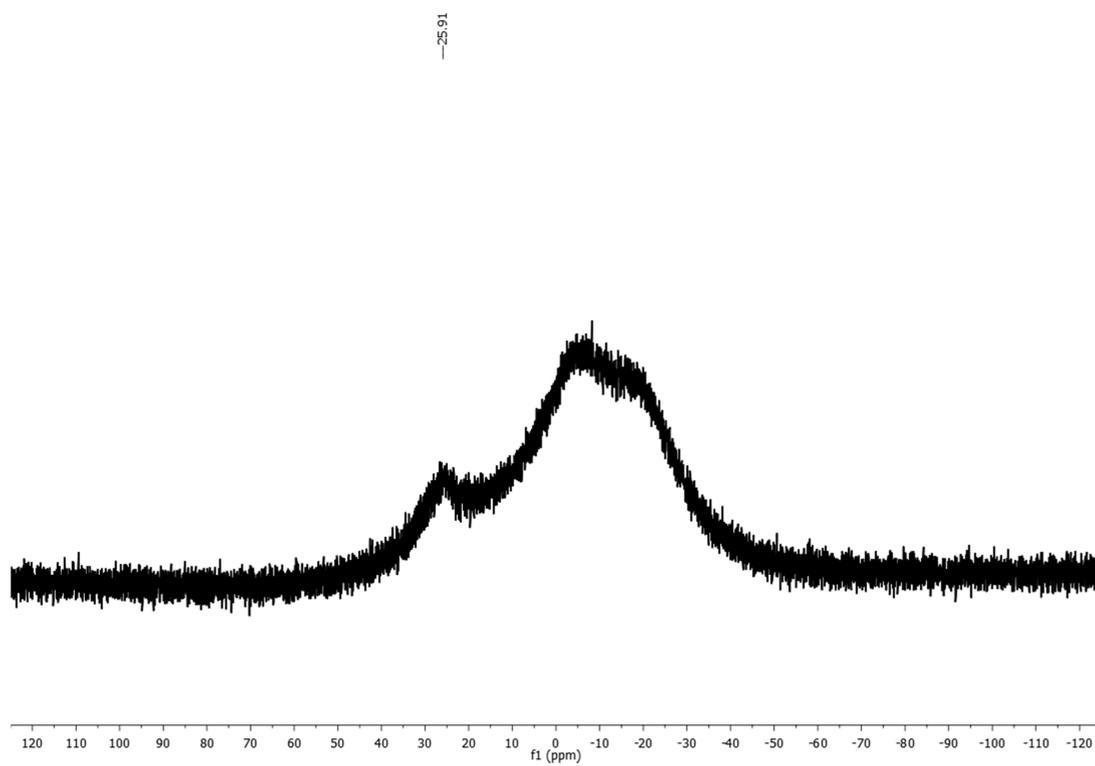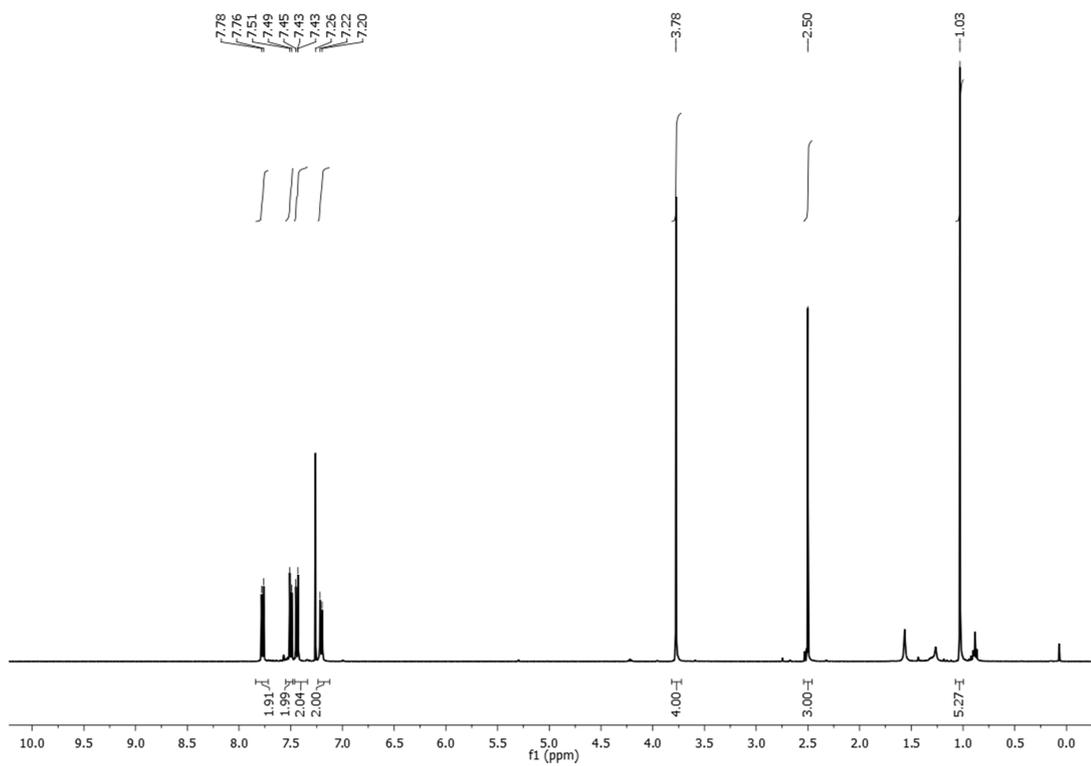
27c – ¹H NMR CDCl₃

27c – ¹³C NMR CDCl₃

27c – ¹¹B NMR CDCl₃

—26.65

20a - ¹H NMR CD₃OD

8.15
7.63
7.54
7.54
7.54
7.53
7.53
7.53
7.50
7.39
7.38
7.38
6.12
6.11

4.68
4.67
4.66
4.31
4.31
4.30
4.30
4.13
4.13
3.88
3.88
3.86
3.76
3.75
3.74

0.97
2.02
4.02
1.05
3.08
0.92
1.03
0.93
0.96
0.99
0.98

10.0   9.5   9.0   8.5   8.0   7.5   7.0   6.5   6.0   5.5   5.0   4.5   4.0   3.5   3.0   2.5   2.0   1.5   1.0   0.5   0.0
f1 (ppm)

20a - ¹³C NMR CD₃OD

158.84
152.19
151.50
135.71
133.20
132.53
129.90
129.58
129.54
124.48
123.80
123.49
118.07
90.99
90.97
89.87
87.26
75.62
72.49
63.41

200   190   180   170   160   150   140   130   120   110   100   90   80   70   60   50   40   30   20   10   0
f1 (ppm)

230
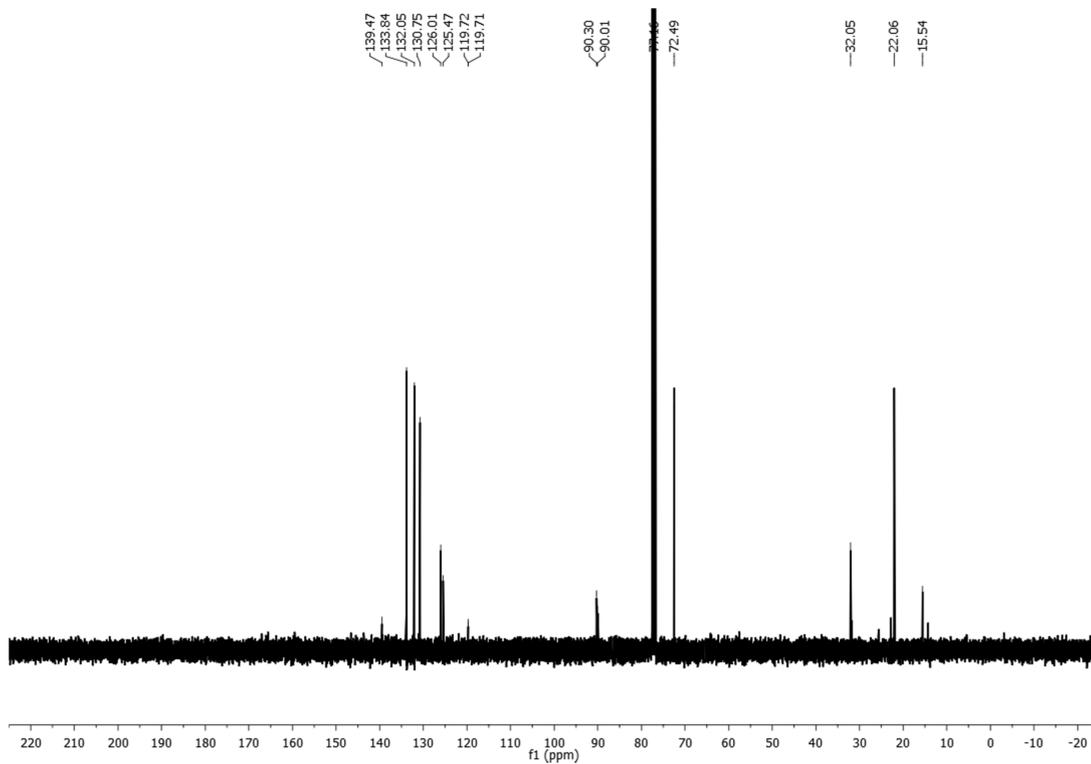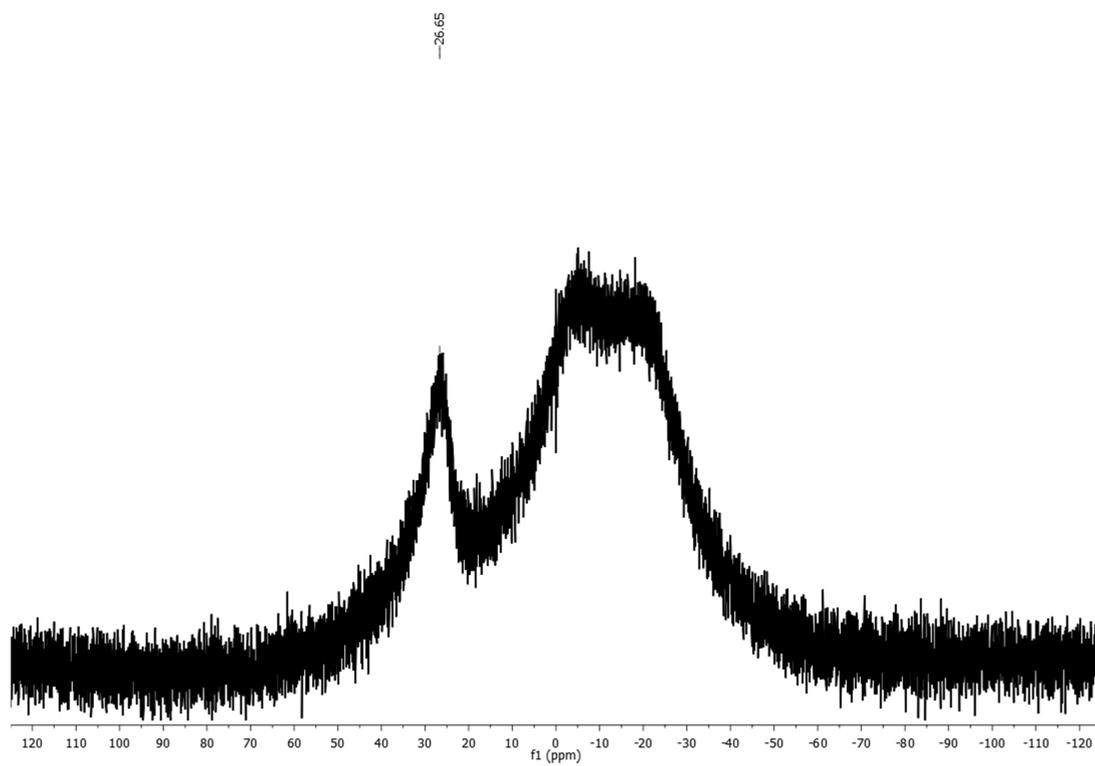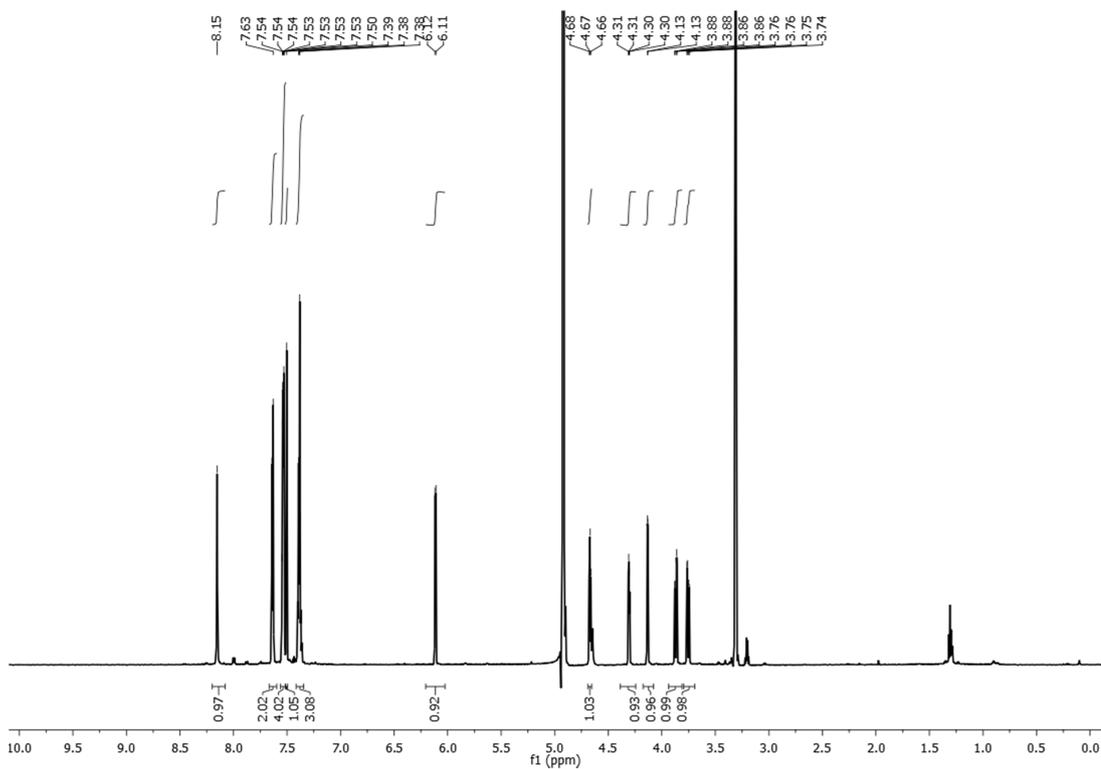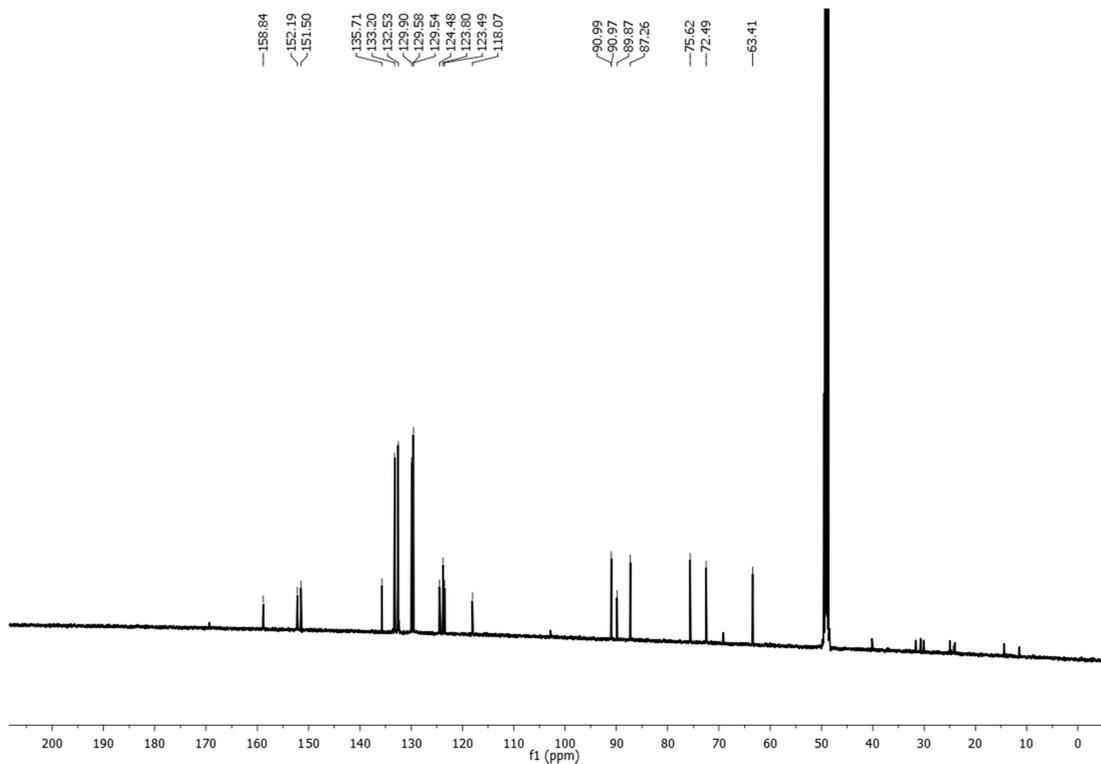
20b – ¹H NMR DMSO-*d*₆

20b – ¹³C NMR CD₃OD

231
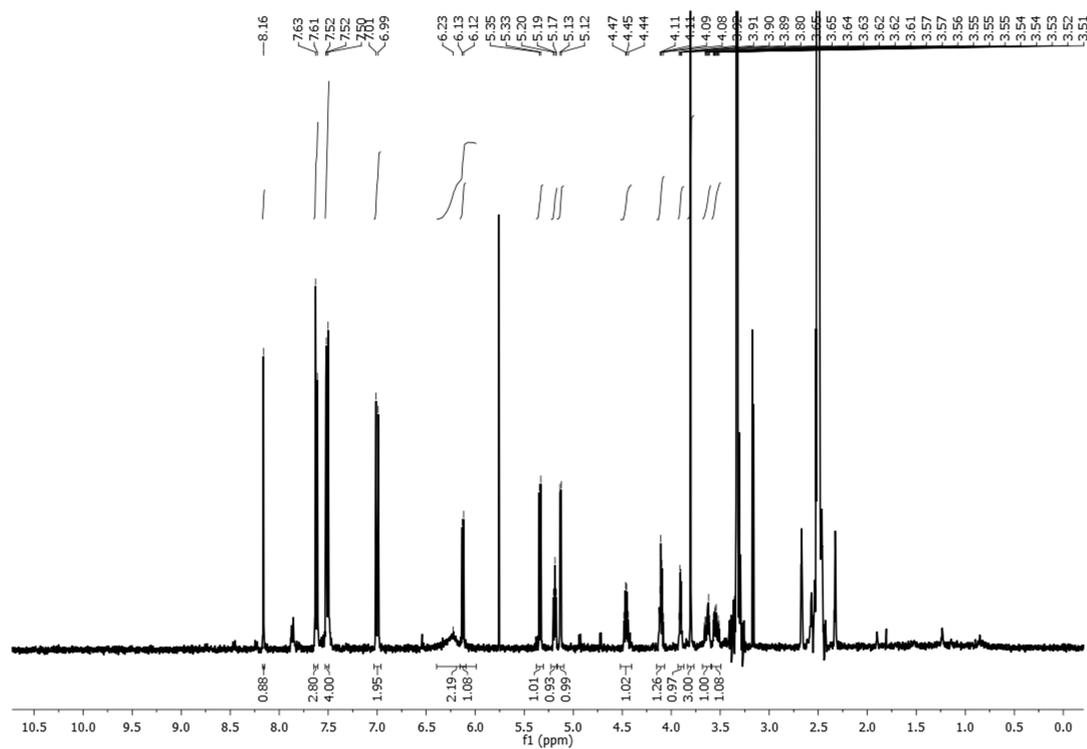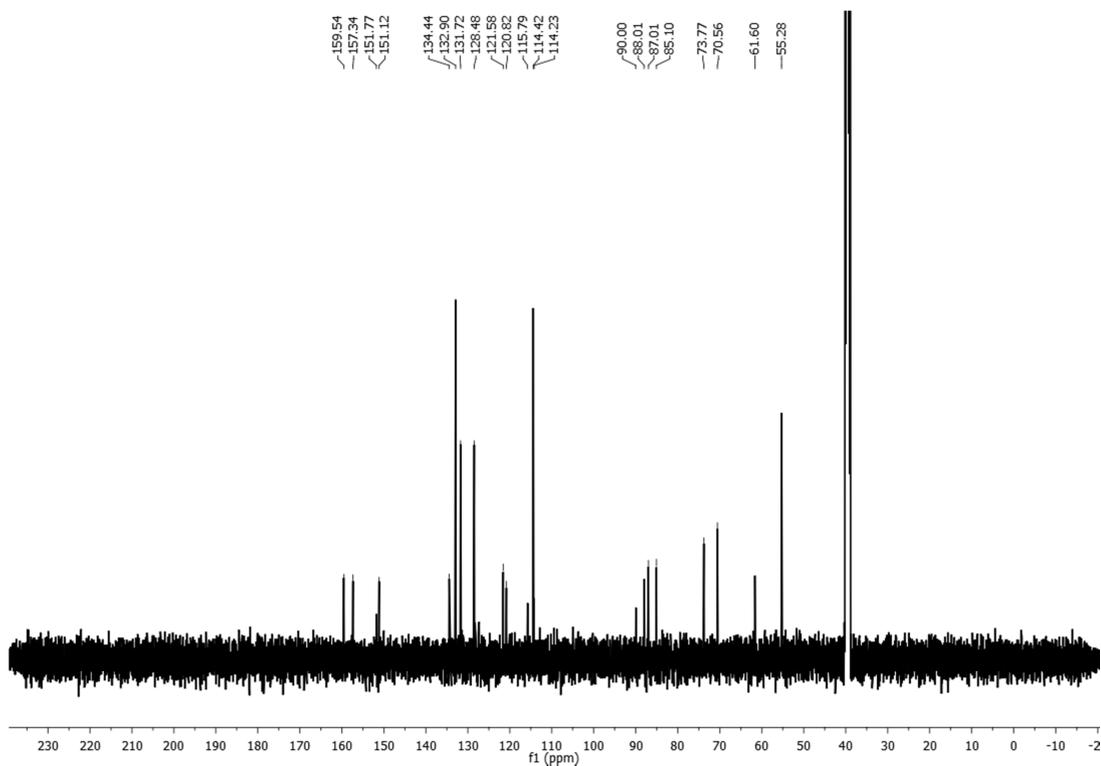
20c – ¹H NMR CD₃OD

20b – HSQC NMR DMSO-*d₆*

232

20d – ¹H NMR CD₃OD

8.15
7.66
7.66
7.65
7.62
7.60
7.54
7.52
7.49
7.46
7.22
6.12
6.10
4.68
4.67
4.65
4.32
4.31
4.30
4.14
4.13
4.13
3.89
3.88
3.86
3.85
3.77
3.76
3.74
3.73

1.07
0.93
1.81
1.75
0.81
0.91
1.00
1.11
1.01
0.91
0.95
1.06
1.08

f1 (ppm)

20d – ¹³C NMR CD₃OD

158.98
152.37
151.57
135.66
133.09
130.72
129.93
129.90
126.86
123.74
123.60
123.48
118.06
91.01
89.24
87.26
86.26
75.62
72.50
63.43
49.85

f1 (ppm)
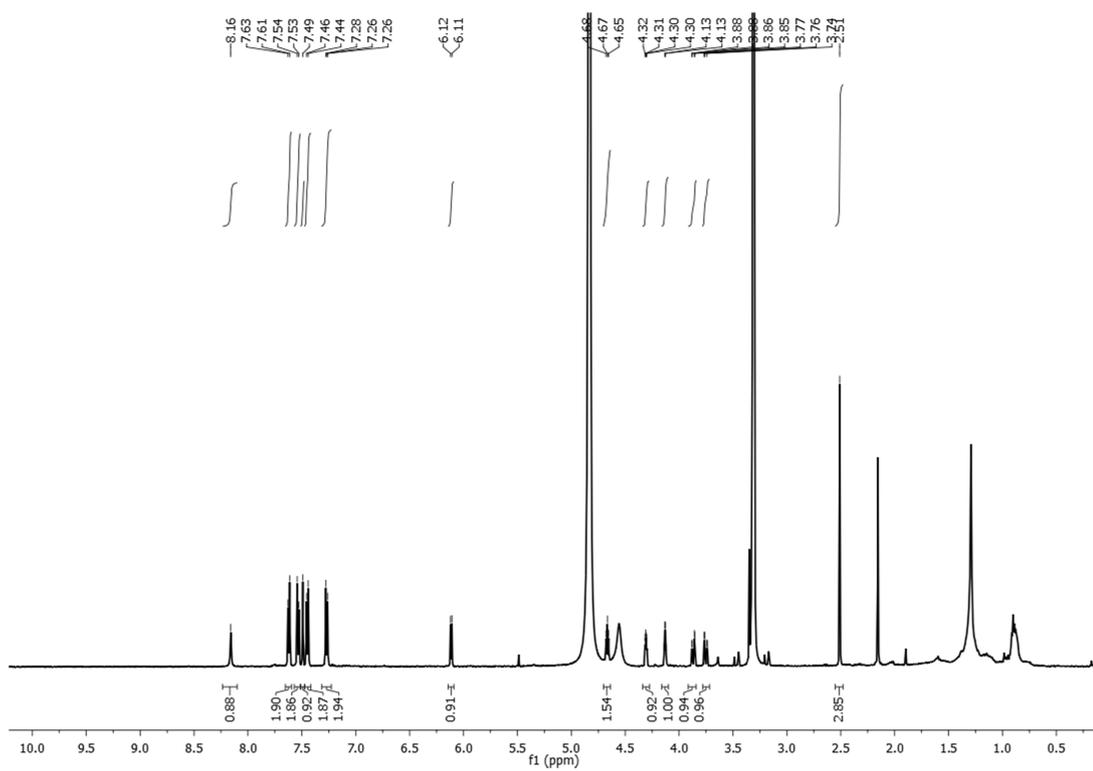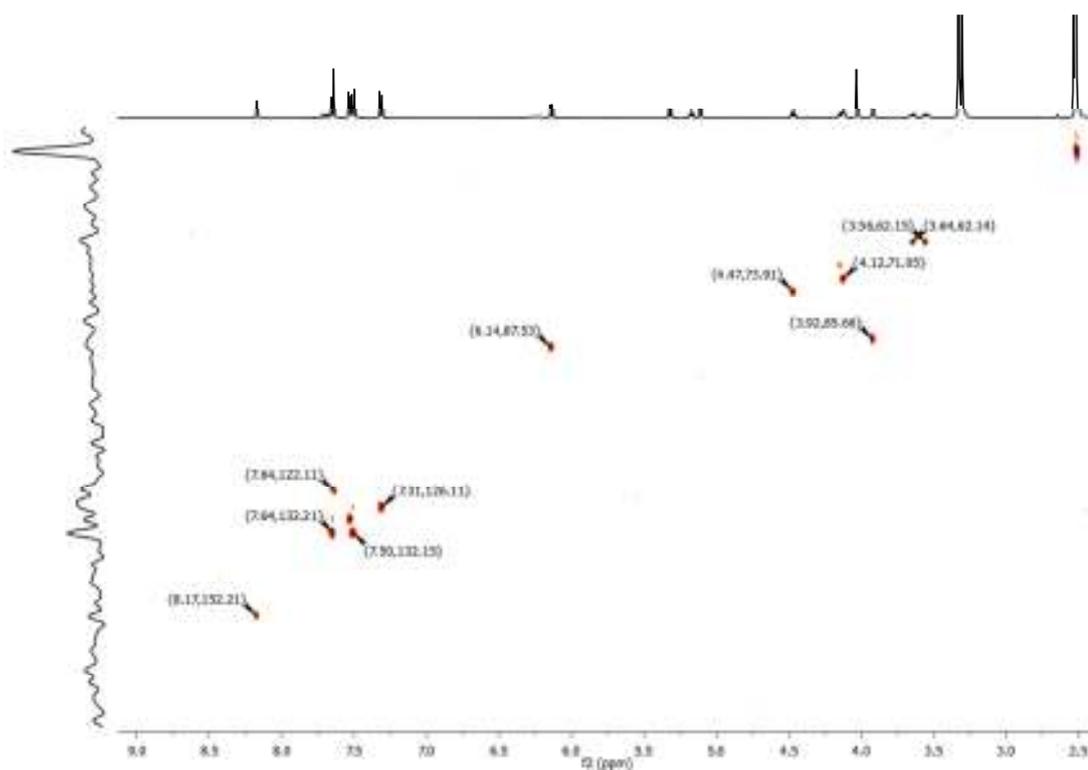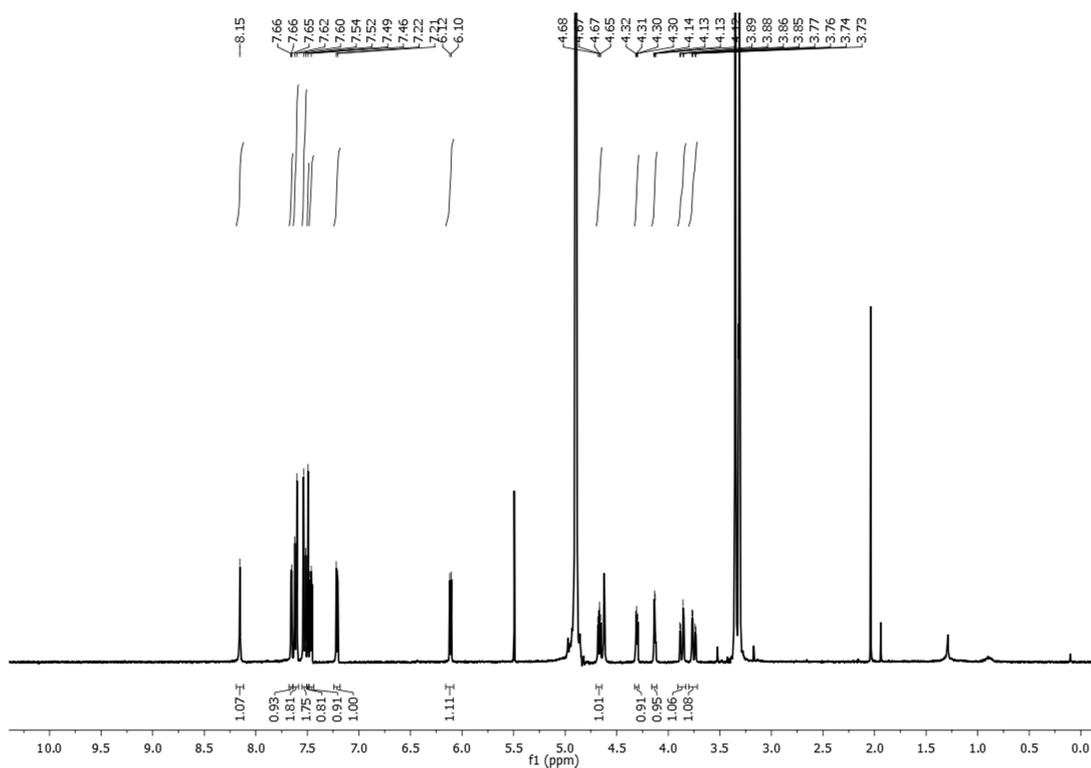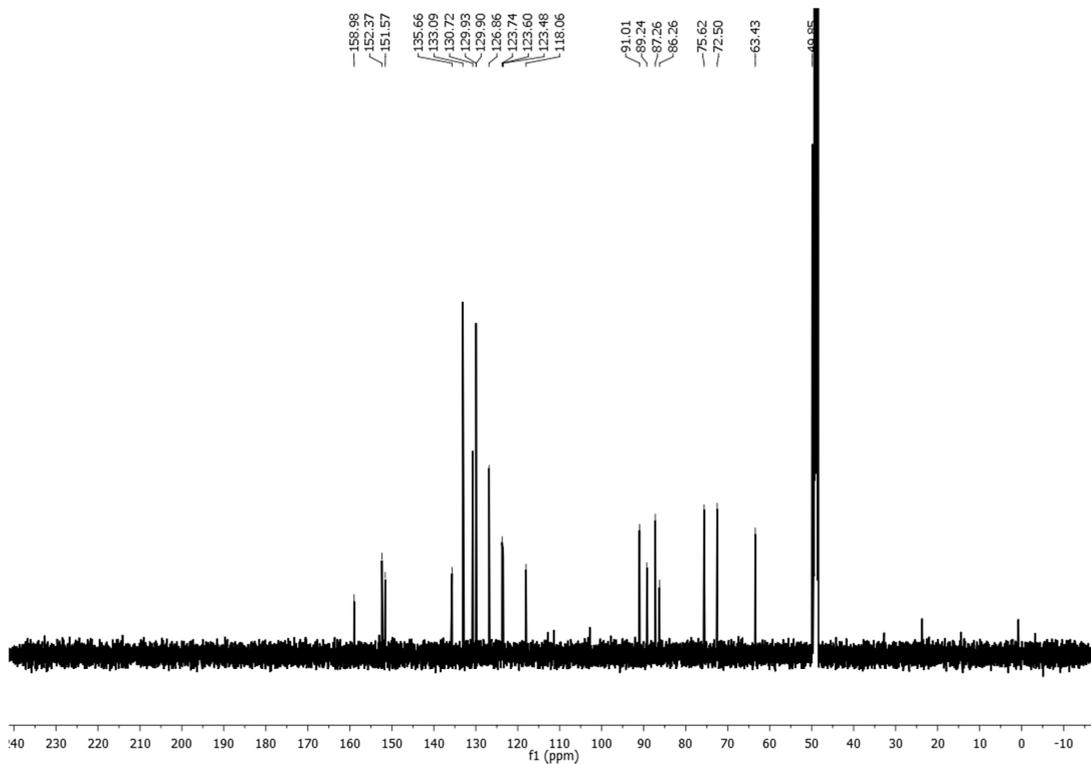
20e – ¹H NMR CD₃OD

20e – ¹³C NMR CD₃OD

234
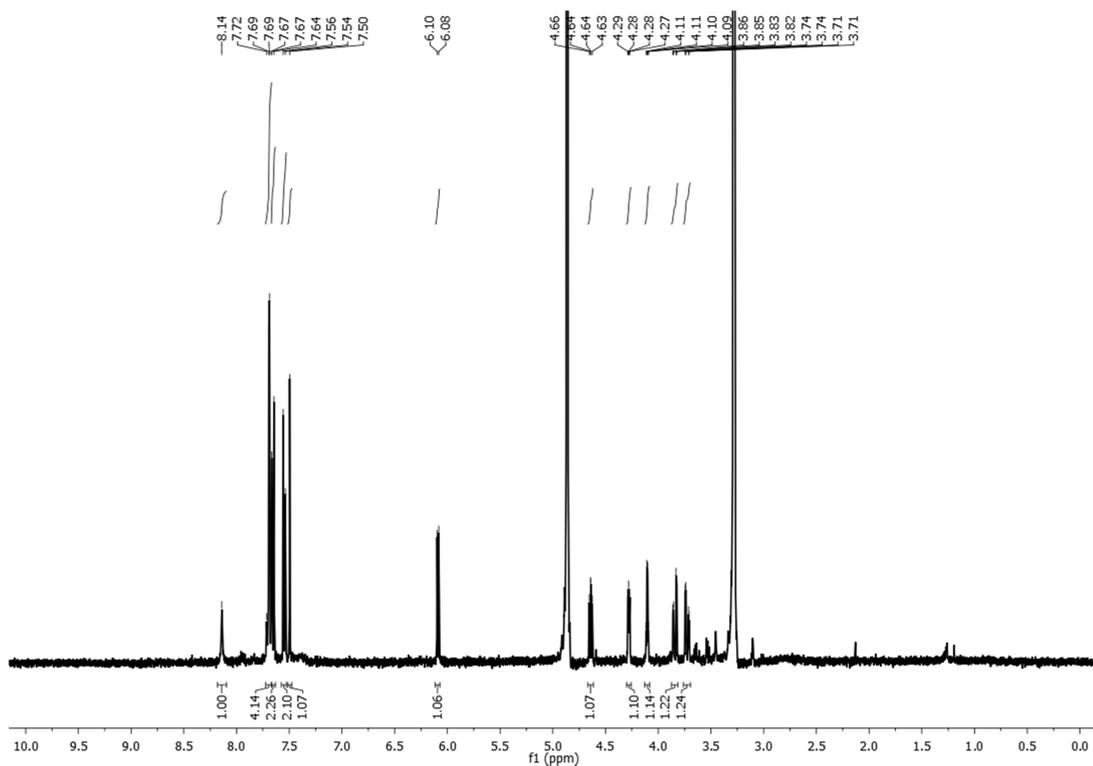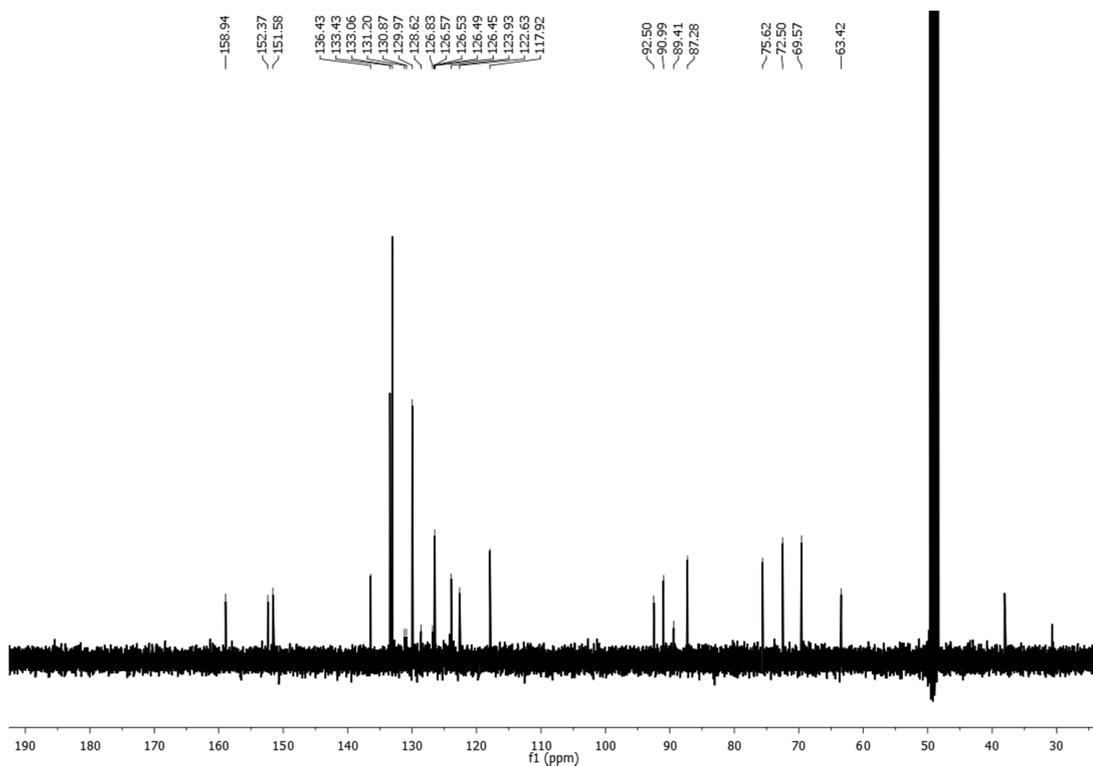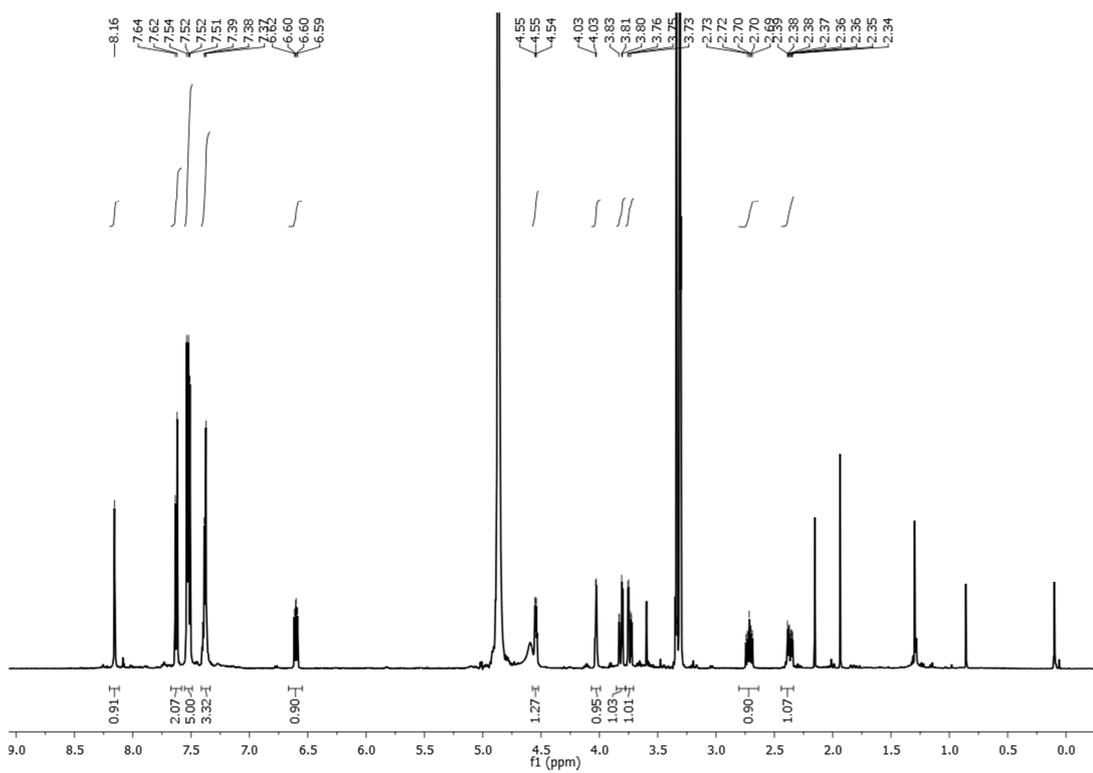
35 – ^{1}H NMR CD$_3$OD

35 – ^{13}C NMR CD$_3$OD
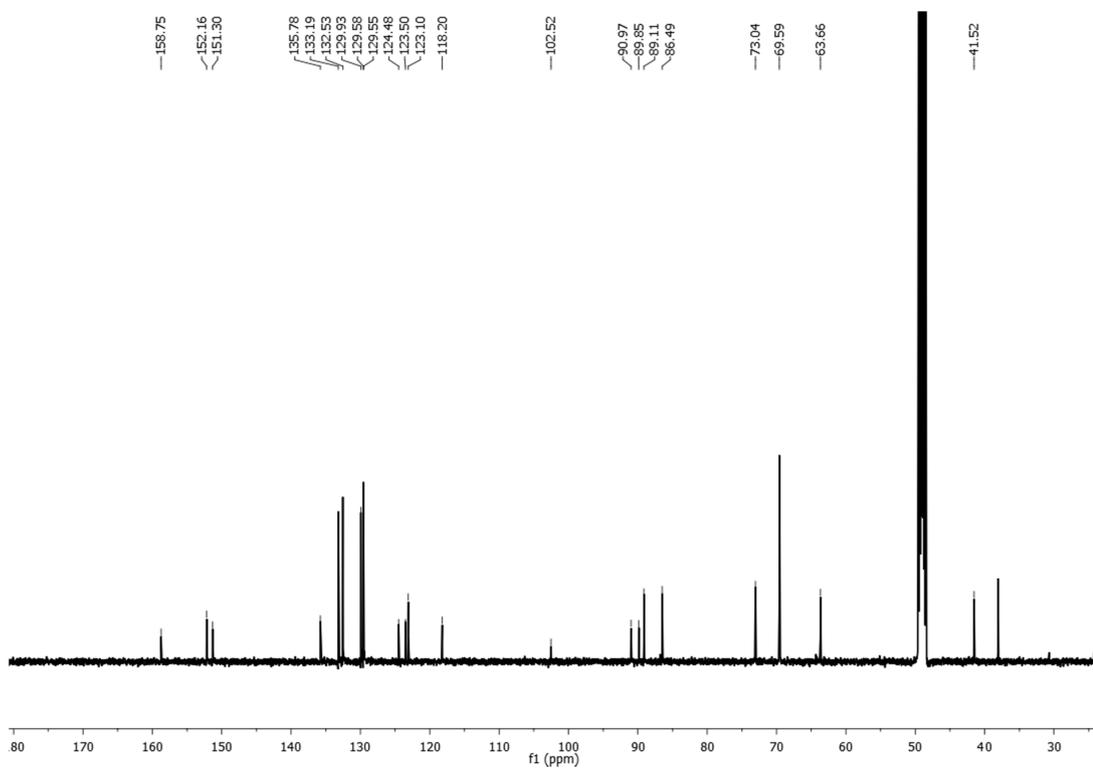
36 – ¹H NMR CD₂Cl₂
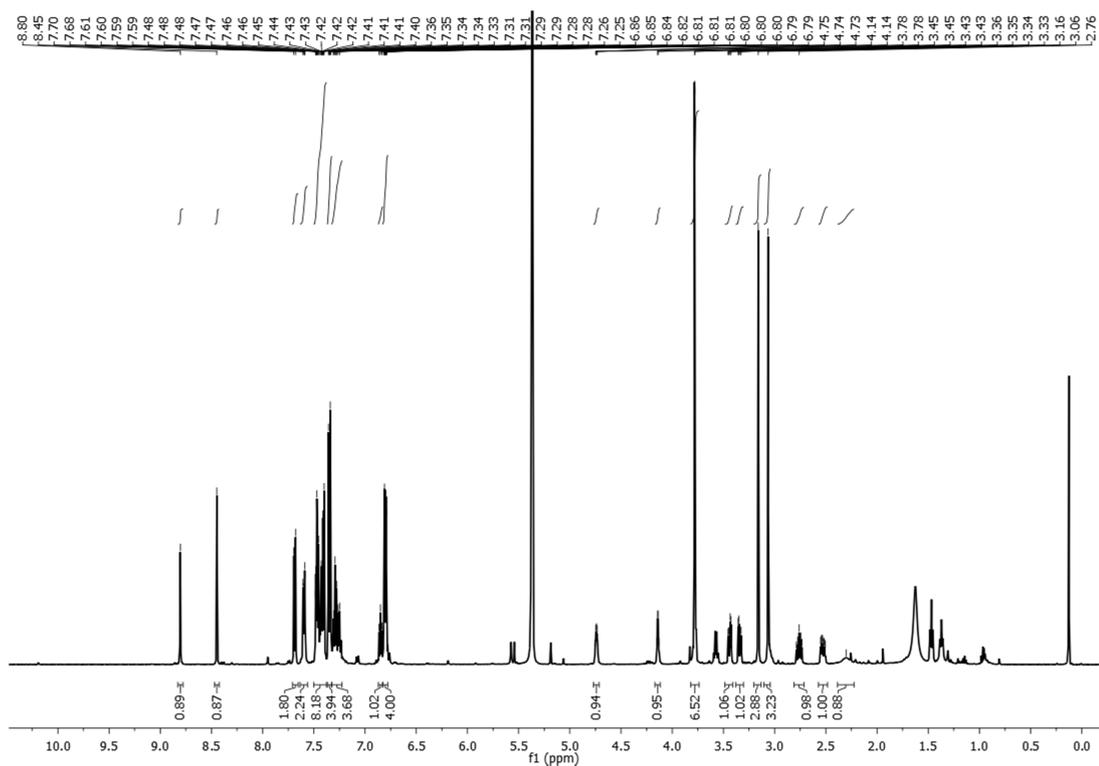
36 – ¹³C NMR CD₂Cl₂
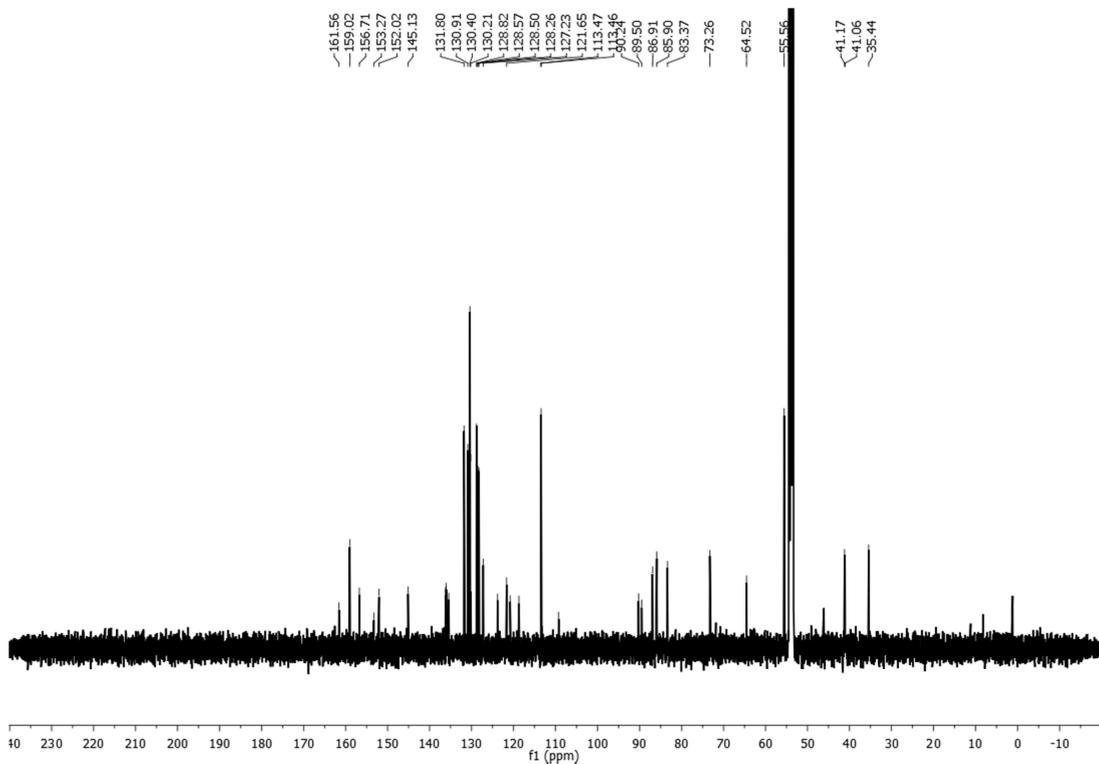
Abbreviations

Ac	acetyl
APCI	atmospheric pressure chemical ionisation
Ar	aryl or heteroaryl
ATP	adenosine triphosphate
bp	boiling point
BSA	(*N*,*O*)-bis(trimethylsilyl)acetamide
Bu	butyl
Bz	benzoyl
c^{7}A	7-deazaadenosine
CD	circular dichroism
COSY	correlation spectroscopy
dba	dibenzylamine acetone
DCE	dichloroethane
DEAE	diethylaminoethyl cellulose
DMA	(*N*,*N*)-dimethylacetamide
DMF	(*N*,*N*)-dimethylformamide
DMF-DMA	(*N*,*N*)-dimethylformamide dimethylacetal
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
dATP	2'-deoxyadenosine triphosphate
dCTP	2'-deoxycytidine triphosphate
dUTP	2'-deoxyuridine triphosphate
EI	electron impact
ESI	electrospray ionization
Et	ethyl
eq	equivalents
FPLC	fast protein liquid chromatography

GTP	guanosine triphosphate
HIV	human immunodeficiency virus
HMDS	hexamethyldisilazane
HOMO	highest occupied molecular orbital
HPLC	high performance liquid chromatography
HRMS	high resolution mass spectrometry
HSQC	heteronuclear single quantum coherence
ICT	intramolecular charge transfer
IMes	1,3-bis(2,4,6-trimethylphenyl)imidazolylidene
IRF	instrument response function
LUMO	lowest unoccupied molecular orbital
Me	methyl
mp	melting point
MS	mass spectrometry
M_w	weight average molecular mass
NAD^+	nicotinamide adenine dinucleotide
n.d.	not determined
NIS	(*N*)-iodosuccinimide
NOESY	nuclear Overhauser effect spectroscopy
NMR	nuclear magnetic resonance
NTP	nucleoside triphosphate
PCR	polymerase chain reaction
Pet Ether 40-60	petroleum ether, bp 40-60 °C
PEX	primer extension
Ph	phenyl
Piv	pivalyl, trimethylacetyl
Pr	propyl

PVP	poly-(*N*)-vinyl-2-pyrrolidone
PXPd	dichlorobis(chlorodi-*tert*-butylphosphine)palladium
RNA	ribonucleic acid
RNAP	ribonucleic acid polymerase
rt	room temperature (*ca.* 12 – 25 °C)
SNP	single nucleotide polymorphism
TBDMS	*tert*-butyldimethylsilyl
TCSPC	time correlated single photon counting
TEAB	triethylammonium bicarbonate
Tf	triflyl, trifluoromethylsulfonyl
TFA	trifluoroacetate
THF	tetrahydrofuran
T_m	melting temperature (denaturation temperature)
TMEDA	(*N,N,N',N'*)-tetramethylethylenediamine
TMS	trimethylsilyl
Tol	*para*-toluoyl
TPPTS	trisodium triphenylphosphine trisulfonate
TRIP	2,4,6-triisopropylphenyl
TXTPS	trisodium tri(4,6-dimethyl-3-sulfonatophenyl)phosphine
UTP	uridine triphosphate
UV-Vis	ultraviolet-visible
XPhos	2-dicyclohexylphosphino-2',4',6'-triisopropylbiphenyl
ε	molar absorption (or extinction) coefficient
λ	wavelength
τ	fluorescence lifetime
Φ	fluorescence quantum yield

References

[1] Watson, J. D.; Crick, F. H. C. *Nature* **1953**, *171*, 737-738.

[2] Stryer, L. *Biochemistry* 3rd ed. 1988, New York: W. H. Freeman and company.

[3] Alberts, B.; Johnson, A.; Lewis, J.; Raff, M.; Roberts, K.; Walter, P. *Molecular Biology of the Cell* 5th ed. 2007, New York: Garland Science.

[4] Saenger, W. *Principles of Nucleic Acid Structure*, 1984, New York: Springer-Verlag.

[5] (i) Jacobsen, K. A.; Jarvis, M. F.; Williams, M. *J. Med. Chem.* **2002**, *45*, 4057-4093, (ii) Škeddelj, V.; Tomašić, T.; Mašič, L. P.; Zega, A. *J. Med. Chem.* **2011**, *54*, 915-929, (iii) McKenna, C. E.; Kashemirov, B. A.; Peterson, L. W.; Goodman, M. F. *Biochemica et Biophysica Acta* **2010**, *1804*, 1223-1230.

[6] Alvarez-Salas, L. M. *Curr. Top. Med. Chem.* **2008**, *8*, 1379-1404.

[7] Coen, D. M.; Richman, D. D. Antiviral Agents, in: *Fields Virology*, Knipe, D. M.; Howley, P. M. Eds. 2007, 447-486, Lippincott Williams & Wilkins. Available from World Wide Web: http://books.google.co.uk/books?id=5O0somr0w18C

[8] Kreutz, C.; Micura, R. Investigations on Fluorine-Labelled Ribonucleic Acids by ^{19}F NMR spectroscopy, in: *Modified Nucleosides in Biochemistry, Biotechnology and Medicine*, Herdewijn, P. Ed.; Wiley-VCH: Weinheim, 2008.

[9] Krstić, I.; Endeward, B.; Margraf, D.; Marko, A.; Prisner, T. F. Structure and Dynamics of Nucleic Acids in:*EPR Spectroscopy: Applications in Chemistry and Biology*, Drescher, M.; Jeschke, G., Eds; Springer: Heidelberg, 2012.

[10] (i) Vanderheiden, B. S. *Anal. Biochem.* **1968**, *22*, 304-310, (ii) Rigby, P. W. J.; Dieckmann, M.; Rhodes, C.; Berg, P. *J. Mol. Biol.* **1977**, *113*, 237-251 and references therein.

[11] (i) Kerman, K.; Kobayashi, M.; Tamiya, E. *Meas. Sci. Technol.* **2004**, *15*, R1-R11, (ii) 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, (iii) Brázdilová, P.; Vrábel, M.; Pohl, R.; Pivoňková, H.; Havran, L.; Hocek, M.; Fojta, M. *Chem. Eur. J.* **2007**, *13*, 9527-9533.

[12] (i) Livak, K. J.; Flood, S. J. A.; Marmaro, J.; Giusti, W.; Deetz, K. *Genome Res.* **1995**, *4*, 357-362, (ii) Kricka, L. J.; Fortina, P. *Clin. Chem.* **2009**, *55*, 670-683, (iii) Yang, Y.; Zhao, L. *Trends Anal. Chem.* **2010**, *29*, 980-1003, (iv) Chudakov, D. M.; Lukyanov, S.; Lukyanov, K. A. *Trends Biotechnol.* **2005**, *23*, 605-613, (v) Marks, G. M.; Nolan, G. P. *Nature Methods* **2006**, *3*, 591-596.

[13] Sinkeldam, R. W.; Greco, N. J.; Tor, Y. *Chem. Rev.* **2010**, *110*, 2579-2619.

[14] (i) Jares-Erijman, E. A.; Jovin, T. M. *Nat. Biotechnol.* **2003**, *21*, 1387-1395, (ii) Ishikawa-Ankerhold, H. C.; Ankerhold, R.; Drummen, G. P. C. *Molecules* **2012**, *17*, 4047-4132, (iii) Preus, S.; Wilhelmsson, L. M. *ChemBioChem* **2012**, *12*, 1990-2001.

[15] Jameson, D. M.; Eccelston, J. F. Fluorescent nucleotide analogues: synthesis and applications, in: *Methods in Enzymology: Fluorescence Spectroscopy* 1997, *278*, 363-390.

[16] Bagshaw, C. R. *J. Cell Sci.* **2001**, *114*, 459-460.

[17] Cremo, C. R. Fluorescence nucleotides: synthesis and characterization, in: *Methods in Enzymology: Biophotonics Pt A* **2003**, *360*, 128-177.

[18] Uesugi, S.; Miki, H.; Ikehara, M.; Iwahashi, H.; Kyogoku, Y. *Tetrahedron Lett.* **1979**, *20*, 4073-4076.

[19] Eid, J.; Fehr, A.; Gray, J.; Luong, K.; Lyle, J.; Otto, G.; Peluso, P.; Rank, D. et al. *Science* **2009**, *323*, 133-138.

[20] Cahová, H.; Pohl, R.; Bednárová, L.; Nováková, K.; Cvačka, J.; Hocek, M. *Org. Biomol. Chem.* **2008**, *6*, 3657-3660.

[21] Anzai, K.; Nakamura, G.; Suzuki, S. *J. Antibiotics* **1957**, *10*, 201-204.

[22] Rao, K. V. *J. Med. Chem.* **1968**, *11*, 939-941.

[23] Nishimura, H.; Katagiri, K.; Sato, K.; Mayama, M.; Shimaoka, N. *J. Antibiotics* **1956**, *9*, 60-62.

[24] Mizuno, Y.; Ikehara, M.; Watanabe, K.; Suzaki, S. *Chem. Pharm. Bull.* **1963**, *11*, 1091-1094.

[25] Smulson, M. E.; Suhadolnik, R. J. *J. Bio. Chem.* **1967**, *242*, 2872-2876.

[26] Yokoyama, S.; Miyazawa, T.; Iitaka, Y.; Yamaizumi, Z.; Kasai, H.; Nishimura, S. *Nature*, **1979**, *282*, 107-109.

[27] Seela, F.; Zulauf, M.; Sauer, M.; Deimel, M. *Helv. Chim. Acta* **2000**, *83*, 910-927.

[28] Kaufmann, G. F.; Meijer, M. M.; Sun, C.; Chen, D.-W.; Kujawa, D. P.; Mee, J. M.; Hoffman, T. Z.; Wirsching, P.; Lerner, R. A.; Janda, K. D. *Angew. Chem. Int. Ed.* **2005**, *44*, 2144-2148.

[29] Jäger, S.; Rasched, G.; Kornreich-Leshem, H.; Engeser, M.; Thum, O.; Famulok, M. *J. Am. Chem. Soc.* **2005**, *127*, 15071-15082.

[30] Ju, J.; Kim, D. H.; Bi, L.; Meng, Q.; Bai, X.; Li, Z.; Li, X.; Marma, M. S.; Shi, S.; Wu, J.; Edwards, J. R.; Romu, A.; Turro, N. J. *Proc. Nat. Acad. Sci.* **2006**, *103*, 19635-19640.

[31] Okamoot, A.; Saito, Y.; Saito, I. *J. Photoch. Photobio. C* **2005**, *6*, 108-122.

[32] Vrábel, M.; Pohl, R.; Votruba, I.; Sajadi, M.; Kovalenko, S. A.; Ernsting, N. P.; Hocek, M. *Org. Biomol. Chem.* **2008**, *6*, 2852-2860.

[33] Brázdilová, P.; Vrábel, M.; Phol, R.; Pivoňková, H.; Havran, L.; Hocek, M.; Fojta, M. *Chem. Eur. J.* **2007**, *13*, 9527-9533.

[34] Cahová, H.; Havran, L.; Brázdilová, P.; Pivoňková, H.; Phol, R.; Fojta, M.; Hocek, M. *Angew. Chem. Int. Ed.* **2008**, *47*, 2059-2062.

[35] Richards, J. B.; Rivadeneira, F.; Inouye, M.; Pastinen, T. M.; Soranzo, N.; Wilson, S. G.; Andrew, T.; Falchi, M.; Gwilliam, R.; Ahmadi, K. R.; Valdes, A. M.; Arp, P.; Whittaker, P.; Verlaan, D. J.; Jhamai, M.; Kumanduri, V.; Moorhouse, M.; van Meurs, J. B.; Hofman, A.; Pols, H. A. P.; Hart, D.; Zhai, G.; Kato, B. S.; Mullin, B. H.; Zhang, F.; Deloukas, P.; Uitterlinden, A. G.; Spector, T. D. *Lancet*, **2008**, *371*, 1505-1512.

[36] Cambien, F.; Poirier, O.; Nicaud, V.; Hermann, S.-M.; Mallet, C.; Ricard, S.; Behague, I.; Hallet, V.; Blanc, H.; Loukaci, V.; Thillet, J.; Evans, A.; Ruidavets, J.B.; Arveiler, D.; Luc, G.; Tiret, L. *Am .J. Hum. Genet.* **1999**, *65*, 183-191.

[37] Halushka, M.; Fan, J.-B.; Bentley, K.; Hsie, L.; Naiping, S.; Weder, A.; Cooper, R.; Lipschutz, R.; Chakravarti, A. *Nature Genet.* **1999**, *22*, 239-247.

[38] Chakravarti, A. *Nature*, **2001**, *409*, 822-823.

[39] (i) Kwok, P.-Y. *Annu. Rev. Genomics Hum. Genet.* **2001**, *2*, 235-258, (ii) Kwok, P.-Y.; Chen, X. *Curr. Issues Mol. Biol.* **2003**, *5*, 43-60.

[40] Olivier, M. *Mut. Res.* **2005**, *573*, 103-110.

[41] Vanderbilt University Sequencing Facility Homepage. http://seq.mc.vanderbilt.edu/DNA (accessed September 2012)

[42] Okamoto, A.; Tainaka, K.; Saito, I. *J. Am. Chem. Soc.* **2003**, *125*, 4972-4973.

[43] Okamoto, A.; Tainaka, K.; Saito, I. *Tetrahedron Lett.* **2003**, *44*, 6871-6874.

[44] Iwasaki, H.; Ota, N.; Nakajima, T.; Shinohara, Y.; Kodaira, M.; Kajita, M.; Emi, M. *J. Hum. Genet.* **2001**, *46*, 32-34.

[45] Okamoto, A.; Tanaka, K.; Fukuta, T; Saito, I. *J. Am. Chem. Soc.* **2003**, *125*, 9296-9297.

[46] Telser, J.; Cruikshank, K.A.; Morrison, L. E.; Netzel, T. L.; Chan, C.-K. *J. Am. Chem. Soc.* **1989**, *111*, 7226-7239.

[47] Kalyanasundaram, K.; Thomas, J. K. *J. Phys. Chem.* **1977**, *81*, 2176-2180.

[48] Okamoto, A.; Kanatani, K.; Saito, I. *J. Am. Chem. Soc.* **2004**, *126*, 4820-4827.

[49] Saito, Y.; Miyauchi, Y.; Okamoto, A.; Saito, I. *Chem. Commun.* **2004**, 1704-1705.

[50] Shinohara, Y.; Matsumoto, K.; Kugenuma, K; Morii, T.; Saito, Y.; Saito, I. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 2817-2820.

[51] Firth, A. G.; Fairlamb, I. J. S.; Darley, K.; Baumann, C. G. *Tetrahedron Lett.* **2006**, *47*, 3529-3533.

[52] Griffin, L. S.; Storr, T. E.; Baumann, C. G.; Fairlamb, I. J. S. *Unpublished results.*

[53] Srivatsan, S. G.; Weizman, H.; Tor, Y. *Org. Biomol. Chem.* **2008**, *6*, 1334-1338.

[54] Xie, Y.; Maxson, T.; Tor, Y. *Org. Biomol. Chem.* **2010**, *8*, 5053-5055.

[55] Cekan, P.; Sigurdsson, S. T. *Chem. Commun.* **2008**, 3393-3395.

[56] Dierckx, A.; Miannay, F.-A.; Gaied, N. B.; Preus, S.; Björck, M.; Brown, T.; Wilhelmsson, L. M. *Chem. Eur. J.* **2012**, *18*, 5987-5997.

[57] Moran, N.; Bassani, D. M.; Desvergne, J.-P.; Keiper, S.; Lowden, P. A. S.; Vyle, J. S.; Tucker, J. H. R. *Chem. Commun.* **2006**, 5003-5005.

[58] Riedl, J.; Pohl, R. Rulíšek, L.; Hocek, M. *J. Org. Chem.* **2012**, *77*, 1026-1044.

[59] (i) Negishi, E.-I. *Acc. Chem. Res.* **1982**, *15*, 340-348, (ii) Kumada, M. *Pure & Appl. Chem.* **1980**, *52*, 669-679.

[60] Stille, J. K. *Angew. Chem. Int. Ed.* **1986**, *25*, 508-524.

[61] Miyaura, N.; Suzuki, A. *Chem. Rev.* **1995**, *95*, 2457-2483.

[62] Tamao, K.; Sumitani, K.; Kumada, M. *J. Am. Chem. Soc.* **1972**, *94*, 4374-4376.

[63] King, A. O.; Okukado, N.; Negishi, E.-I. *J. Chem. Soc., Chem. Commun.* **1977**, 683-684.

[64] Chincilla, R.; Nájera, C. *Chem. Rev.* **2007**, *107*, 874-922.

[65] (i) Luh, T.-Y.; Leung, M.-K.; Wong, K.-T. *Chem Rev.* **2000**, *100*, 3187-3204, (ii) Crudden, C. M.; Glasspoole, B. W.; Lata, C. J. *Chem. Commun.* **2009**, 6704-6716, (iii) Terao, J.; Kambe, N. *Acc. Chem. Res.* **2008**, *41*, 1545-1554.

[66] (i) Agrofoglio, L. A.; Gillaizeau, I.; Saito, Y. *Chem. Rev.* **2003**, *103*, 1875-1916, (ii) The area of Pd-catalysed functionalisation of nucleosides, nucleotides and nucleic acids was recently reviewed in depth for an RSC book chapter: De Ornellas, S.; Williams, T. J.; Baumann, C. G.; Fairlamb, I. J. S. *Palladium-catalysed modification of nucleosides, nucleotides and amino acids* in: C-H and C-X Bond Functionalisation; Ribas, X., Ed. RSC Catalysis Series; RSC: Cambridge, 2013.

[67] Casalnuovo, A. L.; Calabrese, J. C. *J. Am. Chem. Soc.* **1990**, *112*, 4324-4330.

[68] Western, E. C.; Daft, J. R.; Johnson, E. M.; Gannett, P. M.; Shaughnessy, K. H. *J. Org. Chem.* **2003**, *68*, 6767-6774.

[69] Western, E. C.; Shaughnessy, K. H. *J. Org. Chem.* **2005**, *70*, 6378-6388.

[70] Collier, A.; Wagner, G. *Org. Biomol. Chem.* **2006**, *4*, 4526-4532.

[71] Pesnot, T.; Kempter, J.; Schemies, J.; Pergolizzi, G.; Uciechowska, U.; Rumpf, T.; Sippl, W.; Jung, M.; Wagner, G. K. *J. Med. Chem.* **2011**, *54*, 3492-3499.

[72] Capek, P.; Pohl, R.; Hocek, M. *Org. Biomol. Chem.* **2006**, *4*, 2278-2284.

[73] Cahova, H.; Havran, L.; Brazdilova, P.; Pivonkova, H.; Pohl, R.; Fojta, M.; Hocek, M. *Angew. Chem. Int. Ed.* **2008**, *47*, 2059-2062.

[74] Omumi, A.; Beach, D. G.; Baker, M.; Gabryelski, W.; Manderville, R. A. *J. Am. Chem. Soc.* **2011**, *133*, 42-50.

[75] Volpini, R.; Costanzi, S.; Lambertucci, C.; Vittori, S.; Klotz, K.-N; Lorenzen, A.; Cristalli, G. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 1931-1934.

[76] Flasche, W.; Cismas, C.; Herrmann, A.; Liebscher, J. *Synthesis* **2004**, *14*, 2335-2341.

[77] (i) Capek, P.; Cahova, H.; Pohl, R.; Hocek, M.; Gloeckner, C.; Marx, A. *Chem. Eur. J.* **2007**, *13*, 6196-6203, (ii) Kielkowski, P.; Pohl, R.; Hocek, M. *J. Org. Chem.* **2011**, *76*, 3457–3462.

[78] Cho, J. H.; Prickett, C. D.; Shaughnessy, K. H. *Eur. J. Org. Chem.* **2010**, 3678-3683.

[79] Kottysch, T.; Ahlborn, C.; Brotzel, F.; Richert, C. *Chem. Eur. J.* **2004**, *10*, 4017-4028.

[80] (i) Alberico, D.; Scott, M. E.; Lautens, M. *Chem. Rev.* **2007**, *107*, 174-238, (ii) Ackermann, L.; Vincent, R.; Kapdi, A. R. *Angew. Chem. Int. Ed.*, **2009**, *48*, 9792-9826, (iii) Bellina, F.; Rossi, R. *Tetrahedron*, **2009**, *65*, 10269-10310, (iv) Dyker, G. (ed.) *Handbook of C-H Transformations (Vol. 1)*, **2005**, Weinheim: Wiley-VCH.

[81] Čerňa, I.; Phol, R.; Klepetářová, B.; Hocek, M. *Org. Lett.* **2006**, *8*, 5389-5392.

[82] Čerňa, I.; Phol, R.; Hocek, M. *Chem. Commun*. **2007**, 4729-4730.

[83] Storr, T. E.; Firth, A. G.; Wilson, K.; Darley, K.; Baumann, C. G.; Fairlamb, I. J. S. *Tetrahedron* **2008**, *64*, 6125-6137.

[84] Widegren, J. A.; Finke, R. G. *J. Mol. Catal A.* **2003**, *198*, 317-341 and references therein.

[85] Ellis, P. J.; Fairlamb, I. J. S.; Hackett, S. F. J.; Wilson, K.; Lee, A. F. *Angew. Chem. Int. Ed.* **2010**, *49*, 1820-1824.

[86] Storr, T. E.; Baumann, C. G.; Thatcher, R. J.; De Ornellas, S.; Whitwood, A. C.; Fairlamb, I. J. S. *J. Org. Chem.* **2009**, *74*, 5810-5821.

[87] Pivsa-Art, S.; Satoh, T.; Kawamura. Y.; Miura, M.; Nomura, M.; *Bull. Chem. Soc. Jpn.* **1998**, *71*, 467-473.

[88] (i) Bellina, F.; Cauteruccio, S.; Mannina, L.; Rossi, R.; Viel, S. *J. Org. Chem.*, **2005**, *70*, 3997-4005, (ii) Bellina, F.; Cauteruccio, S.; Mannina, L.; Rossi, R.; Viel, S. *Eur. J. Org. Chem.*, **2006**, *3*, 693-703.

[89] Saladino, R.; Crestini, C.; Ciciriello, F.; Di Mauro, E.; Costanzo, G. *J. Biol. Chem.* **2006**, *281*, 5790-5796.

[90] Muzart, J. *Tetrahedron* **2009**, *65*, 8313-8323.

[91] Storr, T. E.; Strohmeier, J. A.; Baumann, C. G.; Fairlamb, I. J. S. *Chem. Commun.* **2010**, *46*, 6470-6472.

[92] Davoll, J. *J. Chem. Soc.* **1960**, 131-138.

[93] Ramasamy, K.; Imamura, N.; Robins, R. K.; Revankar, G. R. *Tetrahedron Lett.* **1987**, *28*, 5107-5110.

[94] Seela, F.; Rosemeyer, H. *Helv. Chim. Acta*, **1988**, *71*, 1573-1585.

[95] Kazimierczuk, Z.; Cottam, H. B.; Revankar, G. R.; Robins, R. K. *J. Am. Chem. Soc.* **1984**, *106*, 6379-6382.

[96] Hoffer, M. *Chem. Ber.* **1960**, *93*, 2777-2781.

[97] Watanabe, K. A.; Hollenberg, D. H.; Fox, J. J. *J. Carbohydrates Nucleosides Nucleotides* **1974**, *1*, 1.

[98] Bookser, B. C.; Raffaele, N. B. *J. Org. Chem.* **2007**, *72*, 173-179.

[99] Vorbrüggen, H.; Höfle, G. *Chem. Ber.* **1980**, *114*, 1256-1268.

[100] Niedballa, U.; Vorbrüggen, H. *J. Org. Chem.* **1974**, *39*, 3654-3660, 3660-3663, 3663-3667, 3668-3671, 3672-3674.

[101] Vorbrüggen, H.; Krolikiewicz, K.; Bennua, B. *Chem. Ber.* **1981**, *114*, 1234-1255.

[102] Tolman, R. L.; Tolman, G. L.; Robins, R. K.; Townsend, L. B. *J. Heterocycl. Chem.* **1970**, *7*, 799, cited in: Seela, F.; Ming, X. *Tetrahedron* **2007**, *63*, 9850-9861.

[103] Seela, F.; Ming, X. *Tetrahedron* **2007**, *63*, 9850-9861.

[104] (i) Bocchi, V.; Palla. G. *Synthesis* **1982**, 1096-1097. (ii) Witulski, B.; Buschmann, N.; Bergsträβer, U. *Tetrahedron* **2000**, *56*, 8473-8480.

[105] Čapek, P.; Cahová, H.; Pohl, R.; Hocek, M.; Gloeckner, C.; Marx, A. *Chem. Eur. J.* **2007**, *13*, 6196-6203.

[106] Hocek, M.; Fojita, M.; *Org. Biomol. Chem.* **2008**, *6*, 2233–2241.

[107] Cho, J. H.; Prickett, C. D.; Shaughnessy, K. H.; *Eur. J. Org. Chem.* **2010**, 3678-3683.

[108] Kim, D.-S.; Ham, J., *Org. Lett.* **2010**, *12*, 1092-1095.

[109] Molander, G. A.; Sandrock, D. L. *J. Am. Chem. Soc.* **2008**, *130*, 15792-15793.

[110] Molander, G. A.; Cavalcanti, L. N. Canturk, B.; Pan, P.-S.; Kennedy, L. E. *J. Org. Chem.* **2009**, *74*, 7364-7369.

[111] Zheng, S.-L.; Lin, N.; Reid, S.; Wang, B. *Tetrahedron*, **2007**, *63*, 5427-5436.

[112] (a) Carrow, B. P.; Hartwig, J. F. *J. Am. Chem. Soc.*, **2011**, *133*, 2116-2119; (b) Amatore, C.; Jutand, A.; Le Duc, G. *Chem. Eur. J.* **2011**, *17*, 2492-2503.

[113] Thompson, W. J.; Jones, J. H.; Lyle, P. A.; Thies, J. E. *J. Org. Chem.* **1988**, *53*, 2052-2055.

[114] Seela, F.; Zulauf, M. *Synthesis* **1996**, 726-730.

[115] Rolland, V.; Kotera, M.; Lhomme, J. *Synth. Comm.* **1997**, *27*, 3505-3511.

[116] Soula, G. *J. Org. Chem.* **1985**, *50*, 3717-3721

[117] Markley, J. L.; Bax, A.; Arata, Y.; Hilbers, C. W.; Kaptein, R.; Sykes, B. D.; Wright, P. E.; Wüthrich, K. *J. Mol. Biol.* **1998** *280*, 933-952, and references therein.

[118] Stolarski, R.; Dudycz, L.; Sugar, D. *Eur. J. Biochem.* **1980**, *108*, 111-121.

[119] Acharya, P.; Chattopadhyaya, J. *J. Org. Chem.* **2002**, *67*, 1852-1865.

[120] Narukulla, R.; Shuker, D. E. G.; Ramesh, V.; Xu, Y.-Z. *Magn. Reson. Chem.* **2008**, *46*, 1-8.

[121] Rosemeyer, H.; Seela, F. *Helv. Chim. Acta* **1988**, *71*, 1573-1585.

[122] Seela, F.; Zulauf, M. *Synthesis*, **1988**, 670-673.

[123] Altona, C.; Sundaralingam, M. *J. Am. Chem. Soc.* **1972**, *94*, 8205-8211.

[124] Reese, C. B. *Org. Biomol. Chem.* **2005**, *3*, 3851-3868.

[125] Beaucage, S. L.; Iyer, R. P. *Tetrahedron* **1992**, *48*, 2223-2311.

[126] Davies, M. J.; Shah, A.; Bruce, I. J. *Chem. Soc. Rev.* **2000**, *29*, 97-107.

[127] Somoza, A. *Chem. Soc. Rev.* **2008**, *37*, 2668-2675.

[128] Seela, F.; Mersmann, K.; Grasby, J. A.; Gait, M. J. *Helv. Chim. Acta* **1993**, *76*, 1809-1820.

[129] Seela. F.; Mersmann, K. *Helv. Chim. Acta* **1993**, *76*, 1435-1449.

[130] Vongsutilers, V.; Daft, J. R.; Shaugnessy, K. H.; Gannett, P. M. *Molecules* **2009**, *14*, 3339-3352.

[131] Burgess, K.; Cook, D. *Chem. Rev.* **2000**, *100*, 2047-2059.

[132] Ludwig, J. *Acta Biochim. Biophys. Acad. Sci. Hung.* **1981**, *16*, 131-133.

[133] Ruth, J. L.; Cheng, Y.-C. *Mol. Pharmacol.* **1981**, *20*, 415-422.

[134] Yoshikawa, M.; Kato, T.; Takenishi, T. *Bull. Chem. Soc. Jpn.* **1969**, *42*, 3505-3508.

[135] A selection of representative examples: (i) Holzberger, B.; Strohmeier, J.; Siegmund,V.; Diederichsen, U.; Marx, A. *Bioorg. Med. Chem. Lett.* **2012**, *22*, 3136-3139, (ii) Maiti, M.; Siegmund, V.; Abramov, M.; Lescrinier, E.; Rosemeyer, H.; Froeyen, M.; Ramaswamy, A.; Ceulemans, A.; Marx, A.; Herdewijn, P. *Chem. Eur. J.* **2012**, *18*, 869-879, (iii) Zaccolo, M.; Williams, D. M.; Brown, D. M.; Gherardi, E. *J. Mol. Biol.* **1996**, *255*, 589-603, (iv) Srivatsan, S. G.; Tor, Y. *J. Am. Chem. Soc.* **2007**, *129*, 2044-2053, (v) Chiaramonte, M.; Moore, C. L.; Kincaid, K.; Kuchta, R. D. *Biochemistry* **2003**, *42*, 10472-10481.

[136] Ikemoto, T.; Haze, A.; Hatano, H.; Kitamoto, Y. *Chem. Pharm. Bull.* **1995**, *43*, 210-215.

[137] Collier, A.; Wagner, G. *Org. Biomol. Chem.* **2006**, *4*, 4526–4532.

[138] Gillerman, I.; Fischer, B. *Nucleosides, Nucleotides and Nucleic Acids* **2010**, *29*, 245-256.

[139] Ludwig, J.; Eckstein, F. *J. Org. Chem.* **1988**, *54*, 631-635.

[140] Caton-Williams, J.; Lin, L.; Smith, M.; Huang, Z. *Chem. Commun.* **2011**, *47*, 8142-8144.

[141] Caton-Williams, J.; Smith, M.; Carrasco, N.; Huang, Z. *Org. Lett.* **2011**, *12*, 4156-4159.

[142] Firth, A. G. PhD Thesis, University of York **2008**.

[143] Braslavsky, S. E. *Pure Appl. Chem.* **2007**, *79*, 293-465.

[144] Jabloński, A. *Nature* **1933**, *131*, 839-840.

[145] Lakowicz, J. R. *Principles of Fluorescence Spectroscopy*, 3rd ed.; Springer: New York, 2006.

[146] (i) Woodwood, R. B. *J. Am. Chem. Soc.* **1941**, *63*, 1123-1126, (ii) Fieser, L. F.; Fieser, M.; Rajagopalan, S. *J. Org. Chem.* **1948**, *13*, 800-806.

[147] (a) Tretiak, S.; Chernyak, V.; Mukamel, S. *J. Phys. Chem. B* **1998**, *102*, 3310-3315; (b) Melinger, J. S.; Pan, Y.; Kleiman, V. D.; Peng, Z.; Davis, B. L.; McMorrow, D.; Lu, M. *J. Am. Chem. Soc.* **2002**, *124*, 12002-12012

[148] Samori, S.; Tojo, S.; Fujitsuka, M.; Ryhding, T.; Fix, A. G.; Armstrong, B. M.; Haley, M. M.; Majima, T. *J. Org. Chem.* **2009**, *74*, 3776-3782.

[149] Samori, S.; Tojo, S.; Fujitsuka, M.; Yang, S.-W.; Elangovan, A.; Ho, T.-I.; Majima, T. *J. Org. Chem.* **2005**, *70*, 6661-6668.

[150] Zhang, K.; Hu, J.; Chan, K. C.; Wong, K. Y.; Yip, J. H. K. *Eur. J. Inorg. Chem.* **2007**, 384–393.

[151] Okubo, J.; Shinozaki, H.; Koitabashi, T.; Yomura, R. *Bull. Chem. Soc. Jpn.* **1998**, *71*, 329-335.

[152] (i) Storr, T. E. PhD Thesis, University of York, 2010; (ii) Firth, A. G. PhD Thesis, University of York, 2008.

[153] Siddle, J. S.; Ward, R. M.; Collings, J. C.; Rutter, S. R.; Porrès, L.; Applegarth, L.; Beeby, A.; Batsanov, A. S.; Thompson, A. L; Howard, J. A. K.; Boucekkine, A.; Costuas, K.; Halet, J.-F.; Marder, T. B. *New J. Chem.* **2007**, *31*, 841-851.

[154] Hughes, G.; Kreher, D.; Wang, C.; Batsanov, A. S.; Bryce, M. R. *Org. Biomol. Chem.* **2004**, *2*, 3363-3367.

[155] Suzuki, K.; Kobayashi, A.; Kaneko, S.; Takehira, K.; Yoshihara, H.; Shiina, Y.; Oishi, S.; Tobita, S. *Phys. Chem. Chem. Phys.* **2009**, *11*, 9850-9860.

[156] Porrès, L.; Holland, A.; Pålsson, L.-O.; Monkman, A. P.; Kemp, C.; Beeby, A. *J. Fluorescence* **2006**, *16*, 267-272.

[157] Bagshaw, C. R. *J. Cell Science*, **2001**, *114*, 459-460.

[158] Reichardt, C. *Angew. Chem. Int. Ed.* **1979**, *18*, 98-110.

[159] Threlfall, T. L. In *UV Spectroscopy: Techniques, instrumentation and data handling*, Techniques in Visible Ultraviolet Spectrometry Series 4, Clark, B. J.; Frost, T.; Russell, M. A., Eds; Chapman&Hall: London, 1993; Chapter 6.

[160] Steigman, A. E.; Miskowski, V. M.; Perry, J. W.; Coulter, D. R. *J. Am. Chem. Soc.* **1987**, *109*, 5884-5886.

[161] Rosemeyer, H.; Seela, F. *Helv. Chim. Acta* **1988**, *71*, 1573-1585.

[162] Seela, F.; Ming, X. *Tetrahedron* **2007**, *63*, 9850-9861.

[163] Cahová, H.; Pohl, R.; Bednárová, L.; Nováková, K.; Cvačka, J.; Hocek, M. *Org. Biomol. Chem.* **2008**, *6*, 3657-3660.

[164] Ward, D. C.; Reich, E. *J. Biol. Chem.* **1969**, *244*, 1228-1237.

[165] Kypr, J.; Kejnovská, I.; Renčiuk, D.; Vorlíčová, M. *Nuc. Acids Res.* **2009**, *37*, 1713-1725.

[166] Eftink, M. R.; Ghiron, C. A. *Anal. Biochem.* **1981**, *114*, 199-227.

[167] Lakowicz, J. R. *Principals of Fluorescence Spectroscopy* 3rd ed. **2006**, New York: Springer; Chapter 8.

[168] Efink, M. R.; Ghiron, C. A. *J. Phys. Chem.* **1976**, *80*, 486-493.

[169] Takahashi, M.; Maraboeuf, F.; Sakai, Y.; Yakushiji, H.; Mishima, M.; Shirakawa, M.; Iwai, S.; Hayakawa, H.; Sekiguchi, M.; Nakabeppu, Y. *J. Mol. Biol.* **2002**, *319*, 129-139.

[170] Sprous, D.; Young, M. A.; Beveridge, D. L. *J. Mol. Biol.* **1999**, *285*, 1623-1632.

[171] Olsthoorn, C. S. M.; Bostelaar, L. J.; De Rooij, J. F. M.; Van Boom, J. H.; Altona, C. *Eur. J. Biochem.* **1981**, *115*, 309-321.

[172] Bloomfield, V. A.; Crothers, D. M.; Tinoco Jr, I. *Nucleic Acids: Structures, Properties and Functions* **2000**, Sausalito: University Science Books; Chapter 8.

[173] Peng, X.; Li, H.; Seela, F. *Nuc. Acids Res.* **2006**, *34*, 5987-6000.

[174] Rosemeyer, H.; Seela, F. *Helv. Chim. Acta* **1988**, *71*, 1573-1585.

[175] Patel, D. J.; Kozlowski, S. A.; Ikuta, S.; Itakura, K. *Biochemistry* **1984**, *23*, 3218-3226.

[176] Kalnik, M. W.; Kouchakdjian, M.; Li, B. F. L.; Swann, P. F.; Patel, D. J. *Biochemistry* **1988**, *27*, 100-108.

[177] Hunter, W. N.; Brown, T.; Anand, N. N.; Kennard, O. *Nature*, **1986**, *320*, 552-555.

[178] Allawi, H. T.; SantaLucia Jr., J. *Biochemistry* **1998**, *37*, 9435-9444.

[179] Puglisi, J. D.; Wyatt, J. R.; Tinoco Jr., I. *Biochemistry* **1990**, *29*, 4215-4226.

[180] Malkov, V. A.; Voloshin, O. N.; Veselkov, A. G.; Rostapshov, V. M.; Jansen, I.; Soyfer, V. N.; Frank-Kamenetskii, M. D. *Nuc. Acids Res.* **1993**, *21*, 105-111.

[181] Huang, J.; Chan, J.; Chen, Y.; Borths, C. J.; Baucom, K. D.; Larsen, R. D.; Faul, M. M. *J. Am. Chem. Soc.*, **2010**, *132*, 3674-3675.

[182] Fors, B. P.; Krattiger, P.; Strieter, E.; Buchwald. S. L. *Org. Lett.* **2008**, *10*, 3505-3508.

[183] Burns, M. J.; Fairlamb, I. J. S. *Unpublished results* 2006-present.

[184] Ellis, P. J.; Fairlamb, I. J. S.; Hackett, S. F. J.; Wilson, K.; Lee, A. F. *Angew. Chem. Int. Ed.* **2010**, *49*, 1820-1824.

[185] Wipf, P. *Synthesis* **1993**, 537-557.

[186] Boutadla, Y.; Davies, D. L.; Macgregor, S. A.; Poblador-Bahamonde, A. I. *Dalton Trans.* **2009**, 5820-5831.

[187] Lapointe, D.; Fagnou, K. *Chem. Lett.* **2010**, *39*, 1118-1126.

[188] Wang, M; Fan, T.; Lin, Z. *Organometallics* **2012**, *31*, 560-569.

[189] (i) Molander, G. A.; Sandrock, D. L. *J. Am. Chem. Soc.* **2008**, *130*, 15792-15793, (ii) Lennox, A. J. J.; Lloyd-Jones, G. C. *J. Am. Chem. Soc.* **2012**, *134*, 7431-7441.

[190] Pangborn, A. B.; Giardello, M. A.; Grubbs, R. H.; Rosen, R. K.; Timmers, F. J. *Organometallics* **1996**, *15*, 1518-1520.

[191] Williams, D. B. G.; Lawton, M. *J. Org. Chem.* **2010**, *75*, 8351-8354.

[192] Fulmer, G. R.; Miller, A. J. M.; Sherden, N. H.; Gottlieb, H. E.; Nudelman, A.; Stoltz, B. M.; Bercaw, J. E.; Goldberg, K. I. *Organometallics*, **2010**, *29*, 2176-2179.

[193] Sun, Q.-Y.; He, J.; Xu, Z.; Huang, G.; Zhou, X.-P.; Zeller, M.; Hunter, A. D. *Chem. Commun.* **2007**, 4779-4781.

[194] Tour, J. M.; Rawlett, A. M.; Kozaki, M.; Yao, Y.; Jagessar, R. C.; Dirk, S. M.; Price, D. W.; Reed, M. A.; Zhou, C.-W.; Chen, J.; Wang, W.; Campbell, I. *Chem. Eur. J.* **2001**, *7*, 5118-5134.

[195] Alfa Aesar catalogue

[196] Yashima, E.; Nimura, T.; Matsushima, T.; Okamoto, Y. *J. Am. Chem. Soc.* **1996**, *118*, 9800-9801.

[197] Adamo, M. F. A; Pergoli, R. *Org. Lett.* **2007**, *9*, 4443-4446.

[198] Rolland, V.; Kotera, M.; Lhomme, J. *Synth. Comm.* **1997**, *27*, 3505-3511.

[199] Seela, F.; Zulauf, M.; Rosemeyer, H.; Reuter, H. *J. Chem Soc. Perkin Trans. 2* **1996**, 2373-2376.

[200] Adapted from a procedure developed by A. G. Firth.

[201] Experimental detail courtesy of R. M. Edkins and A. Beeby.