SYNTHESIS OF TRIPLEX-FORMING OLIGONUCLEOTIDE CONJUGATES OF THE ANTICANCER DRUG TEMODAL

ANDREW JOHN WALSH

Doctor of Philosophy

ASTON UNIVERSITY
March 1999

Aston University

Synthesis of triplex-forming oligonucleotide conjugates of the anticancer drug temodal

A thesis submitted by Andrew Walsh BSc (Hons) for the degree of Doctor of Philosophy

March 1999

Abstract: Covalent attachment of the anticancer drugs temozolomide (Temodal) and mitozolomide to triplex-forming oligonucleotides (TFOs) is a potential way of targeting these alkylating agents to specific gene sequences to maximise site-selectivity. In this work, polypyrimidine TFO conjugates of both drugs were synthesised and targeted to duplex DNA in an attempt to effect site-specific alkylation of guanine residues. Concurrently, in an attempt to enhance the triple helix stability of TFOs at neutral pH, the thermal stabilities of triplexes formed from TFOs containing isoguanine, 2-*O*-benzyl- and 2-*O*-allyl-adenine were evaluated.

A novel cleavage and deprotection procedure was developed which allowed for the solid phase synthesis of the base-sensitive TFO-drug conjugates using a recently developed silyl-linked controlled pore glass (SLCPG) support. Covalent attachment of either temozolomide or mitozolomide at the 5'-end of TFO conjugates caused no destabilisation of the triplexes studied. The synthesis of a phosphoramidite derivative of mitozolomide enabled direct incorporation of this reagent into a model sequence during DNA synthesis. After cleavage and deprotection of the TFO-drug conjugate, the 5'-end mitozolomide residue was found to have decomposed presumably as a result of ring-opening of the tetrazinone ring. The base-sensitive antibacterial and antitumour agent, metronidazole, was also successfully incorporated at the 5'-end of the oligonucleotide d(T_8) using conventional methods.

Two C2-substituted derivatives of 2'-deoxyadenosine containing 2-*O*-benzyl and 2-*O*-allyl groups were synthesised. Hydrogenolysis of the 2-*O*-benzyl analogue provided a useful route, amenable to scale-up, for the synthesis of the rare nucleoside 2'-deoxyisoguanosine (isoG). Both the 2-*O*-allyl and 2-*O*-benzyl derivatives were incorporated into TFO sequences using phosphoramidite methodology. Thermal melting experiments showed that the 2-*O*-allyl and 2-*O*-benzyl groups caused marked destabilistaion of the triple helices studied, in contrast to hexose-DNA duplexes, where aralkyl substituents caused significant stabilisation of duplexes. TFOs containing isoG were synthesised by Pd(0)-catalysed deallylation of 2-*O*-allyl adenine residues. These sequences containing isoG, in its N3- or O2-H tautomeric form, formed triple helices which were equally as stable as those containing adenine.

Keywords: Nucleoside, Imidazotetrazinone, Triplex, SLCPG, Conjugate, Isoguanine.

To the memory of my dad

Acknowledgements

I would like to thank my supervisor Dr. William Fraser for his constant support and guidance throughout the course of this project.

I am grateful to the following for their much appreciated contributions during this project: Dr. Carl Schwalbe for his assistance in conducting molecular orbital calculations, Dr. Yong Feng Wang and Professor Malcolm Stevens for their helpful discussions, Mike Davis, Karen Farrow and Chris Bache for their technical support and assistance.

I would also like to thank Peter Ashton, Dr. Phil Lowe, Claudine Clark and especially Karen Farrow for performing the mass spectral analyses.

I am very grateful to the Cancer Research Campaign for funding this project.

Thanks to my friends in the Department of Pharmaceutical Sciences for making my stay at Aston so enjoyable.

Finally, I would like to thank my wife Margaret, Tess, Eoghan, my mother Kathleen, Shirley, Eugene and Catherine for their support, encouragement and record-breaking patience! Also, a big thanks to Glen for all her help.

Contents

List of Schemes

List of Figures

List of Tables

Chapter 1

Strategies for Synthesis of Oligonucleotides and Drug Conjugates

1.1 Cancer treatment

1.1.1 Chemotherapeutic agents

A vast array of chemotherapeutic agents has been developed during the last 40 years to combat many different forms of cancer. These antitumour agents are commonly classified by their mechanism of action into six general categories: alkylating agents, antimetabolites, hormonal agents, immunotherapeutic agents, vinca alkaloids and miscellaneous. Of these, it is the alkylating agents which are of particular interest in this project.

1.1.2 Alkylating agents

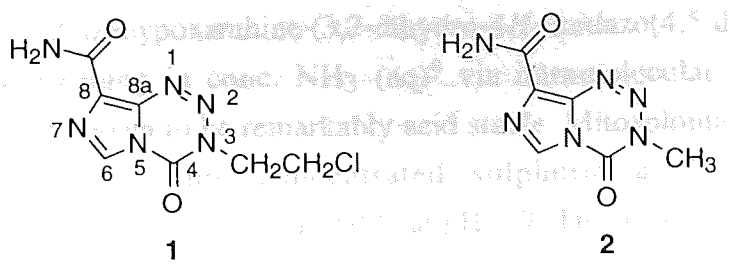
A biological alkylating agent is a compound that can replace a hydrogen atom on a target molecule with an alkyl group under physiological conditions.[1] The precise mechanism by which alkylation leads to cell death remains unclear[2] but it may be that the presence of alkyl groups on DNA play a direct role in cytoxicity by interfering with important protein-DNA interactions and consequently DNA replication. Another possibility is that structural changes such as single-stranded DNA breaks and inter- or intra-strand DNA crosslinks result in cytotoxicity.

There are at least seven major classes of alkylating agents employed in cancer chemotherapy: nitrogen mustards, aziridines, nitrosoureas, triazenes, hydrazine derivatives, methanesulphonic acid esters and imidazotetrazinones.[3, 4, 5] These alkylating agents all react in such a way that an alkyl group or a substituted alkyl group becomes covalently linked to some cellular constituents, usually DNA molecules. Of the seven classes listed above, it is the imidazotetrazinone series of alkylating agents which are the main focus of this thesis work.

1.2 Imidazotetrazinones

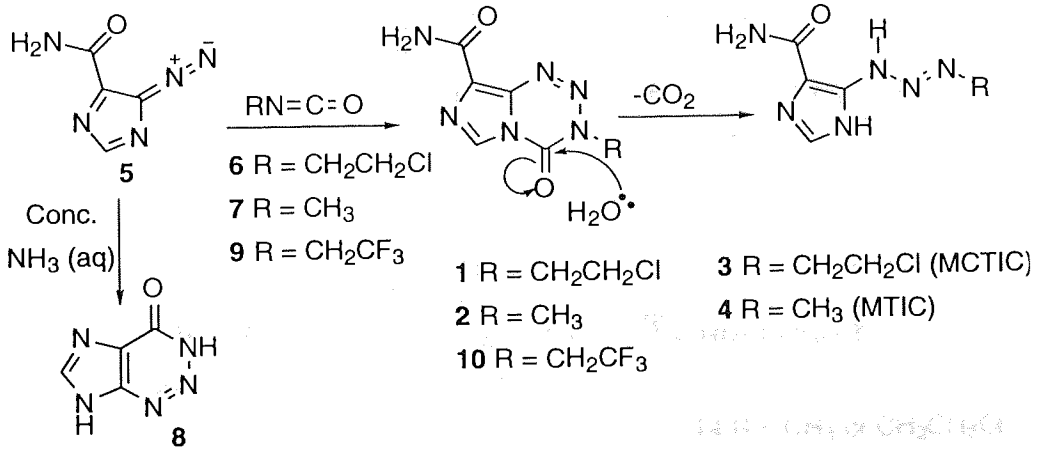
1.2.1 History

The imidazotetrazinones are a fairly recent group of heterocyclic compounds which have been found to exhibit pronounced antitumour effects in a wide range of murine and xenograft tumors.[6] Mitozolomide[7] (**1**), first synthesised in 1980, was the first member of this class to show promising antitumour activity.



1 2

The compound (**1**) is a pro-drug form of the known cytotoxic triazene alkylating agent 5-[3-(2-chloroethyl)triazen-1-yl]imidazole-4-carboxamide (**3**) (MCTIC) (Scheme 1.1). Use of **1** results in alkylation of guanine (G) in DNA, predominantly at the O6 position. The biological activity of mitozolomide has been extensively studied and is considered to exert its effect by crosslinking DNA following O6 alkylation.[8] Both strands of duplex DNA are thus covalently bound together. Temozolomide (**2**), now marketed as Temodal after NDA approval, is a prodrug of the cytotoxic triazene alkylating agent 5-(3-methyltriazen-1-yl)imidazole-4-carboxamide (**4**) (MTIC) (Scheme 1.1) and alkylates the O6 of G. However, cross-linking cannot occur. Conversion of **1** and **2** to **3** and **4** respectively, is achieved by ring-opening of the tetrazinone rings of **1** or **2** by nucleophilic attack of a water molecule at the electron deficient C4 carbonyl which, following spontaneous decarboxylation, yields the triazenes (**3**) and (**4**) respectively (Scheme 1.1). Before considering the precise biological mode of action of **1** and **2**, a brief overview of their synthesis and chemistry is given.



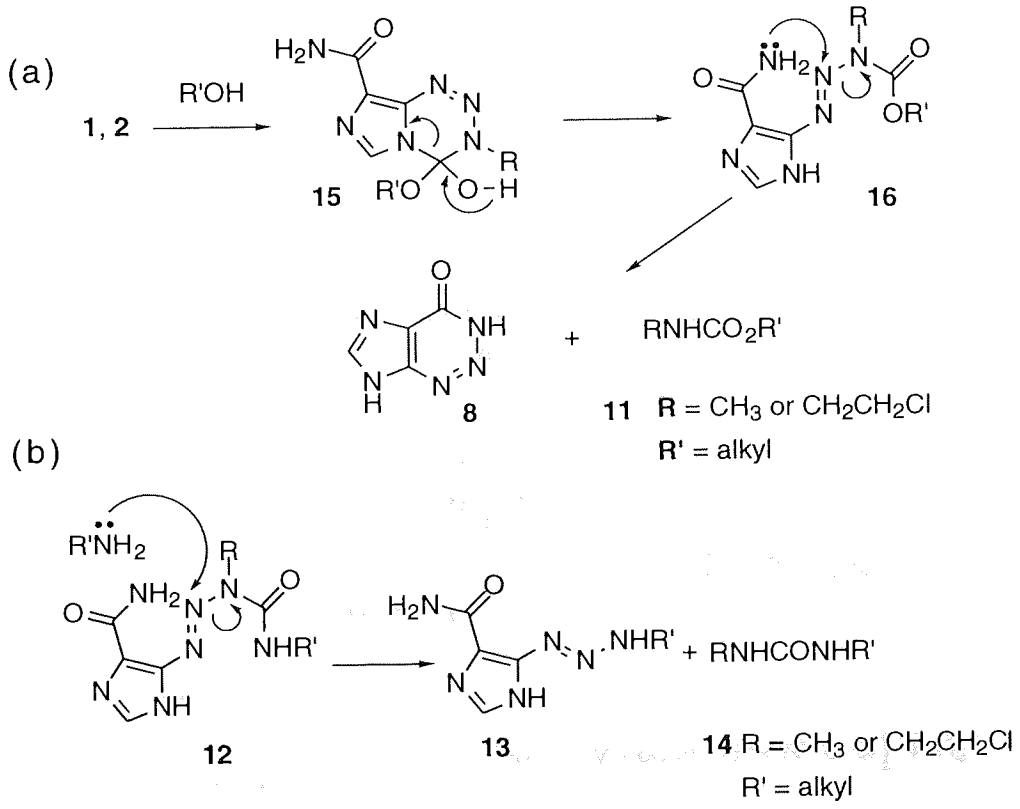
5 8

Scheme 1.1 Reactions of 5-diazoimidazole-4-carboxamide (**5**).

1.2.2 Synthesis and chemical properties of the imidazotetrazinones

Mitozolomide (**1**) and temozolomide (**2**) were first synthesised[7] by reaction of 5-diazoimidazole-4-carboxamide (**5**) with the substituted isocyanates (**6**) and (**7**) respectively (Scheme 1.1). Although **5** is stable to recrystallistion from hot THF, it may be converted

with high efficiency to 2-azahypoxanthine (3,7-dihydro-4*H*-imidazo[4,5-d]-1,2,3-triazin-4-one) (**8**) simply by warming in conc. NH_3 (aq)[9] via intramolecular cyclisation. The imidazotetrazinones are known to be remarkably acid stable. Mitozolomide has even been recovered unchanged from hot concentrated sulphuric acid.[7] However, the imidazotetrazinones are also particularly unstable at pH > 7. This is due to the fact that the C4 carbonyl function is the most electron deficient carbon centre in the molecules (**1**) and (**2**) and thus highly prone to nucleophilic attack. Baig et al[10] studied the reaction of **1** and **2** with a variety of nucleophiles. It was found that in the presence of weak nucleophiles (for example, R'OH wher R' = alkyl) degradation of **1** and **2** followed Scheme 1.2 (a). Nucleophilic attack by R'OH at C4 of **1** or **2** generated the hemiacetal (**15**) which ring opened to give the unstable triazene (**16**) which underwent intramolecular cyclisation generating 2-azahypoxanthine (**8**) and the substituted carbamate (**11**). Baig et al found that addition of conc. NH_3 (aq) led to an acceleration of this reaction involving MeOH as nucleophile.

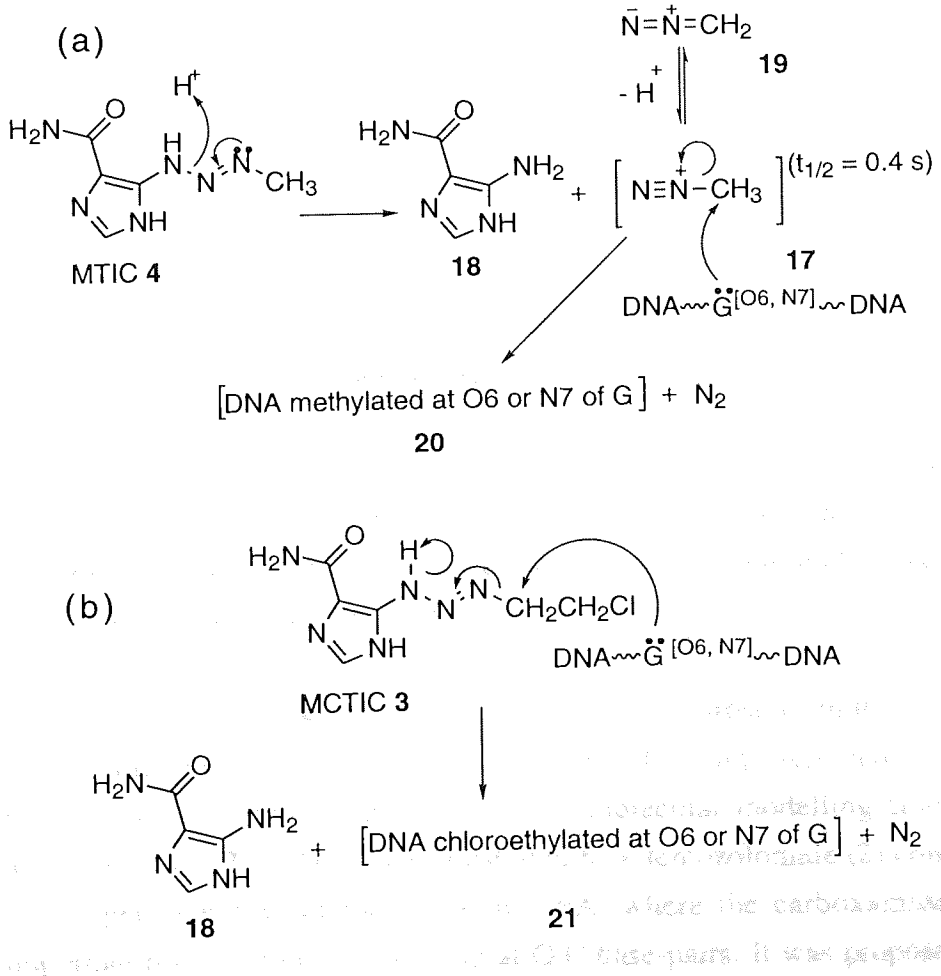


Scheme 1.2[10] Reaction of **1** and **2** with: (a) weak nucleophiles and (b) strong nucleophiles.

Ring opening of the tetrazinone ring-system is always accompanied by loss of the characteristic broad absorption band[7] between 300-350 nm with λ_{max} 325 nm (95% EtOH)

in the UV spectrum of the intact imidazotetrazinones. In the presence of strong nucleophiles ($R'NH_2$), drug decomposition of **1** and **2** follows Scheme 1.2 (b) where, following ring opening, the amine nucleophile intercedes and reacts with the triazene (**12**) forming **13** and urea (**14**).[7, 10] It is of interest to note that retrocyclisation of mitozolomide was achieved by boiling in MeCN[7] or AcOH[10] where product mixtures contained a diazo absorption band (2198 cm^{-1}) in the IR spectrum associated with 5-diazoimidazole-4-carboxamide (**5**). Under milder conditions, it has recently[11] been shown that the trifluoroethyl analogue (**10**) exists in equilibrium with its starting materials (**5**) and (**9**) at 25 °C in anhydrous DMSO (Scheme 1.1) further illustrating the stability of 5-diazoimidazole-4-carboxamide (**5**).

1.2.3 Conversion of the prodrugs 1 and 2 to their bio-active agents



Scheme 1.3 Biological mode of action of: (a) **1** and (b) **2**.

Much work has been carried out to elucidate the mode of action of **2**. It was initially thought that MTIC (**4**) was the DNA alkylating agent, formed as shown in Scheme 1.1.[12]

However, as a result of NMR studies[13, 14] on the decomposition of temozolomide (**2**) and MTIC (**4**) in deuteriated solvent, it has been established that MTIC fragments (Scheme 1.3 (a)) to give the methyl diazonium species (**17**) ($t_{1/2}$ = 0.4 s)[11] as the DNA-methylating agent and 5-aminoimidazole-4-carboxamide (**18**).[14] The methyl diazonium species (**17**) is stabilised by equilibrium with diazomethane (**19**) and reacts with nucleophiles (for example, O6 or N7 of G) generating diatomic nitrogen and the DNA-methylated species (**20**). Mitozolomide (**1**) is converted to the active agent MCTIC (**3**) as outlined in Scheme 1.1. Nucleophilic attack by O6 or N7 of a G base at the chloroethyl group of **3** affords diatomic nitrogen, 5-aminoimidazole-4-carboxamide (AIC, (**18**)) and the chloroethylated species (**21**) (Scheme 1.3 (b)).

1.2.4 Why is guanine specifically alkylated?

The imidazotetrazinones have been shown to alkylate G with a greater preference for the inner G base in a run of three[15, 16, 17] which are prevalent in certain oncogene promoter regions.[18] The reason for the selective alkylation of G is thought to be two-fold:[14]

1. Molecular modelling studies[14] have shown that a sequence of three consecutive guanine bases has a higher dipole moment than sequences of bases chosen from A, T, G and C. This is consistent with the expectation of enhanced nucleophilicity of this major-groove microenvironment and its associated water molecules which might aid ring-opening of the prodrugs (**1**) and (**2**) to the active species (**3**) and (**4**) respectively.

2. Runs of G bases induce localised distortion of the DNA duplex resulting in a wider major groove with the result that the O6 and N7 of G are the most sterically accessible sites in the major groove of DNA. They are also the sites of highest electron density and thus nucleophilicity of all the DNA bases.[19]

Lowe et al[12] proposed a model whereby sequence selectivity arose from the participation of the imidazotetrazinones in an initial noncovalent binding step involving the 8-carboxamide substituent. Denny et al[14] carried out molecular modelling studies of the binding of temozolomide (**2**) to DNA and illustrated how temozolomide (**2**) could make a productive hydrogen bonding interaction with DNA, where the carboxamide function played an important role in orienting the drug at G.C base-pairs. It was proposed that the prodrug was ring opened to its alkylating form by an activated molecule of water in the basic microenvironment of G-rich DNA sequences. Earlier work carried out by Lunt et al[6] demonstrated that a hydrogen bond donor group is preferred at the C8 position of mitozolomide (**1**). This is consistent with the above proposed model for temozolomide (**2**)

which contains a C8 carboxamide group. However, Clark et al[16] have shown that pretreatment of DNA with unreactive isosteres of temozolomide failed to protect DNA from subsequent exposure to temozolomide (**2**), indicating that there is only weak, if any, association of the drug with DNA. It has also been suggested[14, 16] that the ring-opening reaction may in fact occur in free solution under the influence of local pH rather than under catalysis by a target DNA sequence in the major groove. Studies are in progress to determine which of the many species generated from the degradation of temozolomide (**2**) actually interact with, and ultimately alkylate DNA.[17]

1.2.5 Cell death

Alkylation of G can readily occur at the O6 and N7 positions. Although the imidazotetrazinones extensively alkylate the N7 of G, there is no clear evidence that this reaction is responsible for their cellular toxicity.[15,17] N7 alkylation renders the glycosidic bond of the purine nucleotide more labile than that of the parent unmodified purine nucleotide which results in slow spontaneous depurination.[1,19] Enzymatic action at the depurinated site leads to a strand break in the DNA possibly causing cell death. The alkylation of O6 is thought to be the lesion correlated with cytotoxicity[2] where misreading of G as A leads to insertion of T in subsequent replication steps giving a G.C to A.T point mutation.[17,20] The evidence in support of this comes from the fact that tumour lines expressing high levels of the O6-alkylguanine-DNA alkyltransferase repair protein are refactory to treatment by temozolomide (**2**).[14] It is thought that mitozolomide acts by alkylating the O6 of G. An inter-strand cross-link can then occur where ultimately an ethylene group joins the N1 of G and the N3 of the C base on the opposite strand. Obviously, cross-linking cannot occur with temozolomide. It has been suggested that the dose-limiting toxicity of mitozolomide is probably elicited by the DNA cross-linking lesion.[12]

1.2.6 Early clinical trials and side effects of mitozolomide (1) and temozolomide (2)

Mitozolomide (**1**) entered Phase I clinical trials in 1983. It was reported that the dose-limiting toxicity of mitozolomide was thrombocytopenia (damage to platelets).[12] Recommendations were made for a reduction in dose levels. During a number of Phase II clinical trials in 1987, evaluation was abandoned due to the severe and unpredictable myelosuppression (bone marrow loss) that was encountered despite dosage reduction.[8]

Temozolomide (**2**) was selected for clinical trials on the basis that it displayed clear toxicological and mechanistic advantages over mitozolomide.[14] Phase I clinical trials of temozolomide showed that the drug "only occassionally" exhibited the unpredictable

myleosuppression exhibited by mitozolomide.[8] Compound (**2**) is currently used to treat patients suffering from brain tumours and malignant melanoma.[21] It was the aim of this project to improve site-specificity of both **1** and **2** by synthesisng triplex-forming oligonucleotide (TFO) conjugates of these agents to eliminate or at least reduce the harmful side-effects associated with the parent drugs.

1.3 Improving the specificity of alkylating agents: literature precedents

Alkylating agents have been linked to a variety of molecules including carbohydrates, amino acids, steroids, liposomes and purines/pyrimidines in order to obtain greater tumour specificity through exploitation of the propensity of such molecules to accumulate in tumour cells.[4]

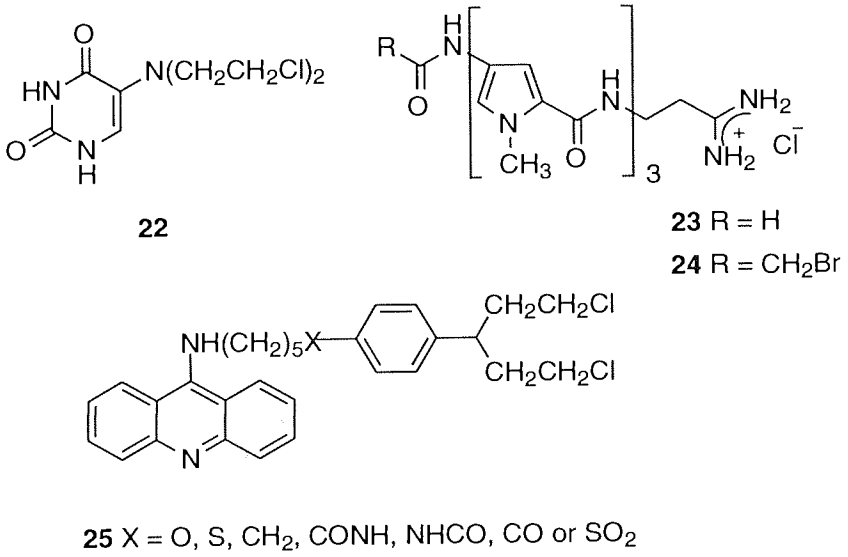


Figure 1.1 Alkylating agents linked to carrier molecules.

For instance, based upon the hypothesis that rapidly dividing cancer cells have a greater requirement for nucleotides than normal cells, purine and pyrimidines containing alkylating functionalities have been synthesised, of which uracil mustard (**22**) is an example (Figure 1.1).[4] Although **22** was found to have good activity, serious side effects including toxicity and carcinogenicity, prevented its clinical use.

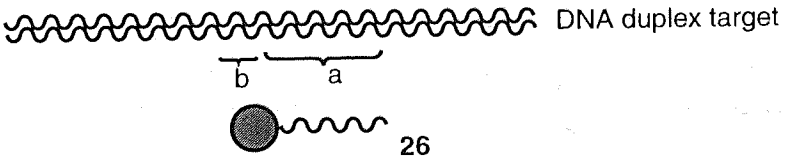
Since alkylating agents actually interact with DNA, the linking of such agents to molecules such as groove binders, intercalating agents and oligonucleotides which are specific for DNA sequences has proved popular. Groove binders, such as the tripeptide unit (**23**) of distamycin, interact sequence specifically with DNA. Oligopyrrole (**23**) binds to five-base-pair sites in the minor groove with a preference for A.T rich regions. Baker et al[22] found

that the bromoacetyl derivative (**24**) (Figure 1.1) selectively alkylated the N3 of adenine adjacent to the binding site.

Intercalating agents have a strong reversible affinity for DNA. They tend to bind DNA in a nonselective manner at almost any sequence.[23] Generally, the planar aromatic chromophore of the intercalator is inserted between the DNA base-pairs. The aniline mustard derivatives (**25**) of 9-amino acridine (Figure 1.1) have proved to be more potent than unconjugated alkylating agents.[24]

It is many years since Belikova[25] suggested that oligonucleotides had potential chemotherapeutic applications in directing DNA reactive functions. However, the discovery in 1987 that short pyrimidine oligonucleotides could bind specifically to duplex DNA targets forming triple helices led to much interest in the application such triplex-forming oligonucleotides (TFOs) as therapeutic agents.[26, 27] In so-called antigene therapy, the DNA double helix can be considered to be a receptor. Once the desired target is known, the sequence, which is specific for that target, is automatically known. Since the base sequence of a 17-mer oligonucleotide occurs just once in the sequence of the human genome[28], antigene oligonucleotides of this length have just a single target in a cell, making this a potentially powerful technique for influencing gene expression by triplex formation.[29] The sequence specific nature of antigene therapy made it an appealing approach for enhancing the site-specificity of the drugs (**1**) and (**2**). Examples of the targeting of antigene TFO-conjugates to DNA sequences are given in Section 1.9.

1.4 Aims



a = Purine duplex target
b = G-rich target site at which drug is known to alkylate
26 = oligonucleotide-drug conjugate

Figure 1.2 Targeting of duplex DNA with a TFO-drug conjugate (**26**).

It was the aim of this project to attach the drugs temozolomide and mitozolomide to antigene TFOs which could be used to target the drugs site-specifically to desired segments of duplex DNA, particularly oncogenic cells. It was anticipated that this antigene approach

would lead to a reduction in the undesirable side-effects (Section 1.2.6) caused by use of the free drugs alone. It was also thought that the use of oligonucleotides to target the drugs to desired sequences would provide a new insight into the mechanism of action of the free drugs. This approach necessitated the use of TFO sequences which could form stable triple helices under physiological conditions.

1.5 The antigene strategy

1.5.1 Double helical DNA

In 1953, Watson and Crick deduced the double helical nature of DNA.[30] It was found that the individual oligonucleotide strands in duplex DNA are held together by the specific base-pairing of A with T and G with C, by means of the Watson-Crick (W-C) hydrogen-bonding patterns shown in Figure 1.3. The G.C base pair (3 hydrogen-bonds (H-bonds)) is more stable than the A.T base pair (2 H-bonds). Factors[23] other than hydrogen-bonding are also known to affect the stability of the DNA duplex. Base stacking is as important a factor as hydrogen bonding in stabilising the duplex. Bases stack when π-σ eletcrostatic attraction outweighs the π-π electron repulsion leading to a reduction in the overall free energy of the system. However, the process of duplex formation leads to a restriction in the free rotation of many bonds thus making the process entropically unfavourable.[23] Duplex formation is therefore enthalpy driven.

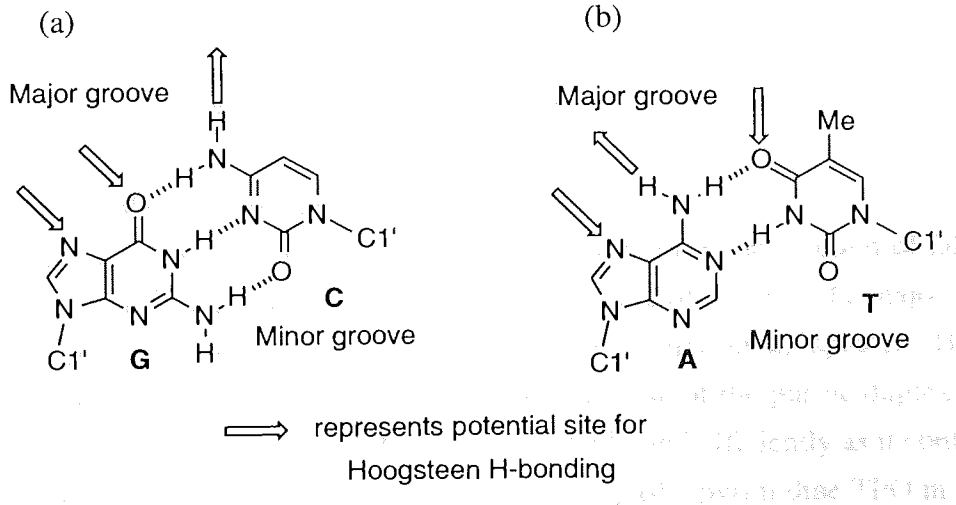


Figure 1.3 Watson-Crick hydrogen bonding between: (a) G.C base pair and (b) A.T base pair.

1.5.2 How is duplex stability assessed?

A common technique for assessing duplex stability is the thermal denaturation experiment[31] which is carried out using a thermostatted UV spectrophotometer. This

technique relies on the fact that duplex formation is accompanied by a hypochromic shift at 260 nm. Thus, by raising the temperature while monitoring the UV absorbance of the duplex, an increase in absorbance may be observed when the W-C hydrogen bonds of the duplex are broken and the bases unstack at a given temperature.

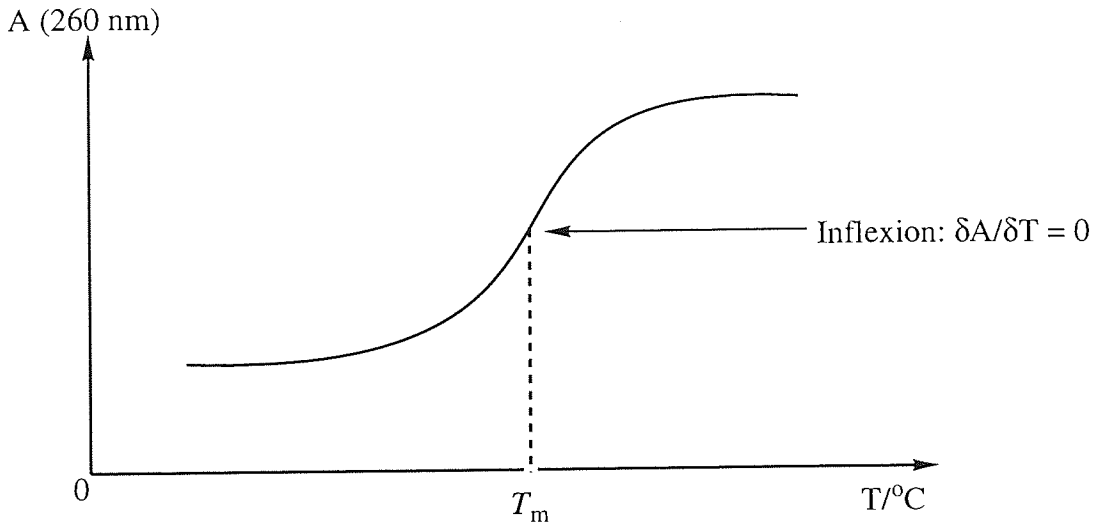


Figure 1.4 Typical thermal melting profile of duplex DNA.

This temperature is called the melting temperature (T_m) of the double helix. Generally, melting occurs across a temperature range. A typical melting curve with A vs T, is shown in Figure 1.4. The T_m is the inflexion point on the curve, that is to say, the point at which the first derivative (δA/δT) of the melting curve equals zero.

1.5.3 DNA triple helix formation

Within a few years of the discovery of duplex DNA, the first description of DNA triple helices was given.[32] Triplexes are formed by the binding of TFOs in the major groove of duplex DNA. Binding occurs as a result of the formation of specific Hoogsteen hydrogen bonds between the third strand bases and those of the purine duplex strand.[33] The pyrimidine strand of the DNA duplex cannot be bound efficiently as it contains only one vacant H-bonding site (Figure 1.3).[34] The binding of a pyrimidine TFO in the major groove of a duplex is usually accompanied by hypochromism at 260 nm.[35] The thermal stability of triple helices may be determined using thermal denaturation as for duplexes, except that there are usually two transitions in the melting profile, the lower T_m usually corresponding to the triplex-duplex transition while the higher T_m usually corresponds to the duplex-random coil transition.

The gel retardation assay is another technique used to determine triplex formation where gel electrophoresis is used to separate DNA sequences according to size. Charged DNA molecules move through the gel under the influence of an electric field, with larger molecules migrating slower than smaller ones.[36] Generally, the 5'-end of the target purine strand of the duplex is radiolabelled with ^{32}P and thus is the only sequence which may be visualised by autoradiography or by exposure to photographic film during a few hours. When this sequence forms a duplex with its complementary strand under non-denaturing conditions (that is, conditions where the duplex is stable), its gel mobility is retarded due to the greater size of the duplex structure compared with the single strands. When the duplex is mixed with a suitable TFO, the gel mobility of the triplex is retarded compared with that of the duplex, thus providing evidence of triplexation.

The footprinting experiment[36, 37, 38] is another technique used to determine the formation of a triple helix. Here again, the 5'-end of the purine target strand of the duplex is labelled with ^{32}P. The enzyme DNAse I digests DNA sequences by cutting DNA at all locations. Each of the labelled fragments have the 5'-^{32}P label in common while the 3'-end varies in length. Upon exposure of the gel of a digested sequence, an individual band is seen for fragments of differing length. The action of DNAse I is inhibited along any stretch of duplex DNA bound by a TFO indicated by a blank region ('footprint') in the gel across the triplex region. Such experiments may be carried out quantitatively to determine the concentration at the which the TFO binds the duplex target. For the purposes of this work, thermal denaturation was the method of choice since very little sample pretreatment is required with this technique. Also, micromolar quantities of purified oligonucleotide may be analysed by this sensitive technique.

1.5.4 Triplexation: literature precedents

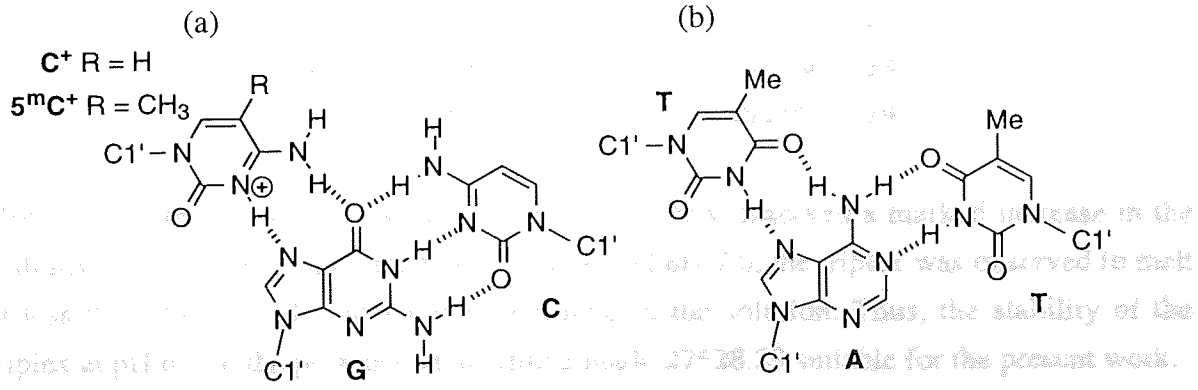


Figure 1.5 Binding schemes for the base triplets: (a) C+*G.C, ^{5m}C+*G.C and (b) T*A.T.

Early work on the formation of triple helices was based upon the specific binding of oligopyrimidine TFOs to the oligopurine sequences of double-helical DNA.[27,39] Specificity was derived from thymine recognition of the A.T base pair (T*A.T base triplet) and protonated cytosine (C^+) recognition of the G.C base pair (C^+*G.C base triplet) (Figure 1.5). The stability of triplexes containing cytosine (C) in the Hoogsteen-paired third strand are stongly pH dependent. This is due to the requirement for the protonation of cytosine N3. Thus, the binding of an antigene oligonucleotide containing C^+ and T to duplex DNA requires conditions considerably more acidic than the intracellular pH range of 7.1-7.6.[40] An important feature of the C^+*G.C and T*A.T triplets is that they are isomorphous. That is to say, the glycosidic, C1'-N bonds of the triplex bases in Figure 1.5 are superimposable which means that the phosphate backbone of the third strand is uniform resulting in a stable triplex. Stability is presumably derived from good stacking of the third strand bases. Thus, a C and T containing TFO-drug conjugate was considered a suitable model sequence for initial studies on targeting of the drugs to duplex DNA.

Many attempts have been made to overcome the pH dependence of triplexation and these are considered further in Chapter 3. A major aim of this work was to synthesise TFO sequences which could form triple helices with improved stability at physiological pH compared with many of the literature precedents. The synthesis and evaluation of novel purine bases for this purpose is described in Chapter 4.

For successful targeting of the drugs to duplex DNA using a TFO, it was necessary that the triple helix formed was stable at body temperature (37 °C). Xiang et al[41] used TFO (**27**) to target the purine strand (**28**) of the duplex (**28.29**) under various conditions. At pH 6.4 they found that the triplex melted at 35.3 °C.

5'-d(TTTCTTTTCTCTCTT) **27**
5'-d(GCGCGAAAGAAAAGAGAGAACCCGG) **28**
3'-d(CGCGCTTTCTTTTCTCTCTTGGGCC) **29**

However, in the presence of spermine (0.5 mM), they observed a marked increase in the stability of the triplex which melted at 45.1 °C. At pH 7.0, the triplex was observed to melt at less than 33 °C with and without spermine in the solution. Thus, the stability of the triplex at pH 6.4 in the presence of spermine made **27*28.29** suitable for the present work.

For the present work, the reported sequences (**27-29**)[41] were modified to read as in triplex **27*30.31** where target G for alkylation is indicated in bold. Incidentally, the use of 5-

methylcytosine (5^mC) in place of C in TFO (**27**) in the presence of spermine resulted in a triplex with a T_m of 42.6 °C at pH 7.0 and 37.9 °C at pH 7.5. Thus, 5^mC was also considered for this thesis work to enable targeting of the drugs to duplex DNA at physiological pH.

5'-d(TTTCTTTTCTCTCTT)	**27**
5'-d(GGG**G**GAAAGAAAAGAGAGAA)	**30**
3'-d(CCCCCTTTCTTTTCTCTCTT)	**31**

1.6 Solid-phase oligonucleotide synthesis

1.6.1 The synthesis cycle

Oligonucleotide synthesis can be performed directly on automated synthesisers using a derivatised solid support of Controlled-Pore-Glass (CPG) and appropriate phosphoramidites. The benzoyl group is commonly used as a protecting group for the exocyclic amino groups of A and C while the isobutyryl group is used to protect the exocyclic amino groups of G. The CPG is derivatised with an appropriate nucleoside attached to the support via a succinate linkage. The general oligonucleotide synthesis cycle is shown in Scheme 1.4.

The standard procedure for solid phase DNA synthesis involves:[31]

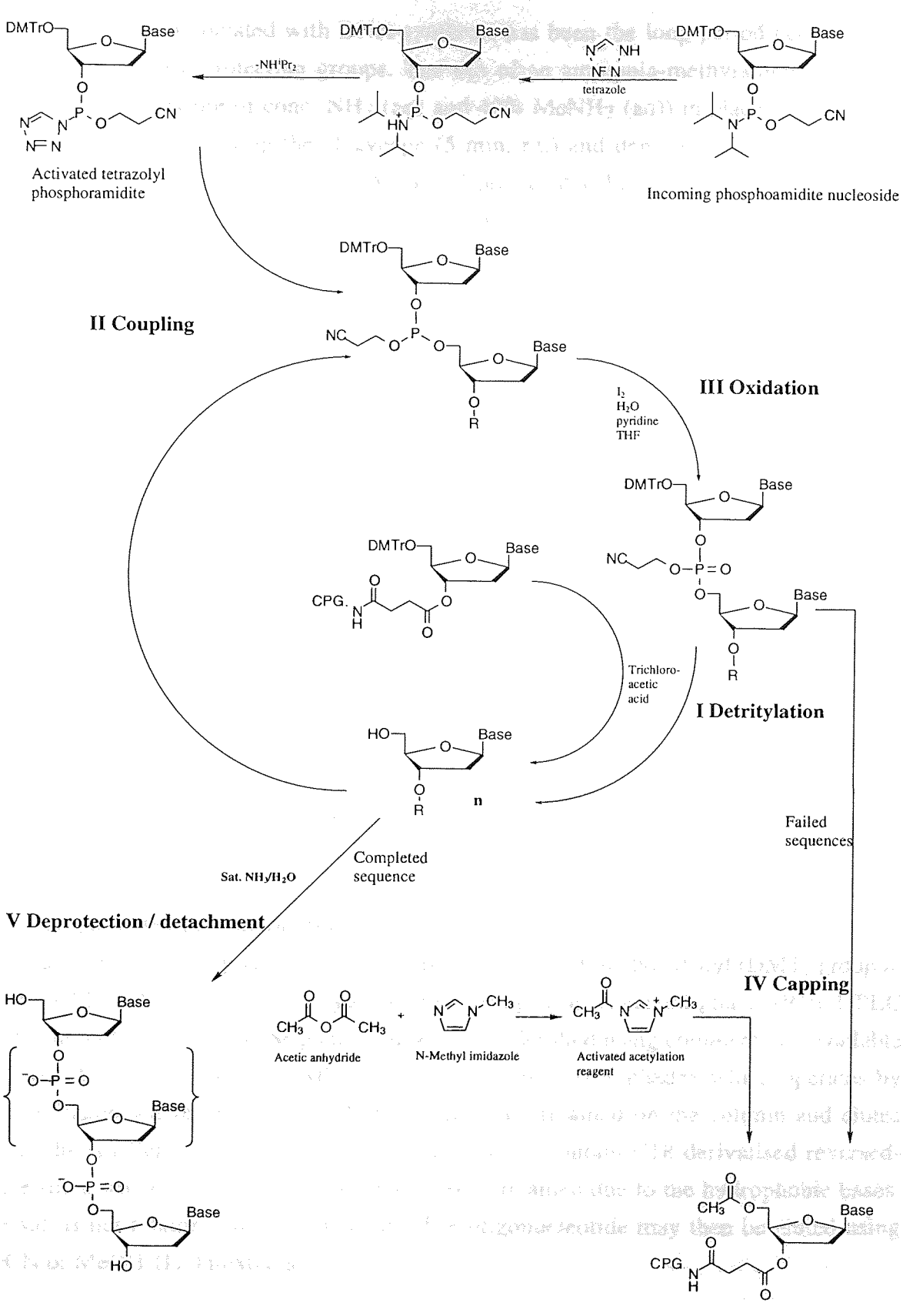
1. Detritylation of the 5'-OH of the support-bound nucleoside using Cl_2CHCO_2H.

2. Activation of the incoming phosphoramidite using tetrazole, which both protonates the phosphoramidite nitrogen and displaces the diisopropylamino group to form the corresponding tetrazolides which are more susceptible to attack by the free 5'-OH.

3. Capping of the 5'-OH groups, by acetylation, of any sequences which failed to couple with the incoming phosphoramidite.

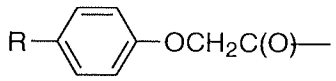
4. Oxidation of phosphite to phosphate linkage using I_2 and H_2O in THF.

5. Cleavage and deprotection of the support-bound oligonucleotide using NH_3 (aq) at 55 °C during 16 h.



Scheme 1.4 The solid-phase DNA synthesis cycle by the phosphoramidite method.

A major problem associated with DNA synthesis has been the long period necessary for removal of the base protecting groups. The use of an ammonia-methylamine[42] (AMA) solution (50:50 mixture of conc. NH_3 (aq) and 40% $MeNH_2$ (aq)) in place of conc. NH_3 (aq) has greatly speeded up the cleaveage (5 min, r.t.) and deprotection (65 °C, 5 min) procedure. A proviso here is that C is *N*4-acetyl-protected as literature reports[43, 44] suggest that, following reaction of benzoyl-protected C with alkylamines, significant amounts (4-20%) of the unwanted *N*4-alkyl derivative is produced. The use of so-called, base-labile amino-protecting groups such as phenoxyacetyl[45] (PAC)) (**32**) or *t*-butylphenoxyacetyl[46, 47] (BPA^t) (**33**) (Figure 1.6)) on A, C and G has allowed cleavage and deprotection of sequences using conc. NH_3 (aq) under more mild conditions compared with the AMA procedure (r.t. during 2 h[48] for PAC and 55 °C during 15 min[46] or r.t. during 2 h[42] for BPA^t).



32 PAC R = H
33 BPA^t R = Bu^t

Figure 1.6 Labile exocyclic amino-protecting groups **32** (PAC) and **33** (BPA^t).

It is to be noted here that the imidazotetrazinones are unstable at pH > 7 and so all conditions described thus far, for cleavage and deprotection of oligonucleotides, are incompatible with **1** and **2**.

1.6.2 Oligonucleotide purification

Sequences are commonly synthesised with the 5'-terminal dimethoxytrityl (DMT) group in place (DMT-ON mode). This hydrophobic group aids reversed-phase (RP) HPLC purification of the sequence. Sequences may then be desalted using commercially available NAP or SEP-PAK columns. NAP columns contain porous Sephadex which operates by size exclusion where the large DNA molecule is not retained on the column and elutes before the smaller salt molecules. SEP-PAK columns contain C18-derivatised reversed-phase silica which cause the oligonucleotide to be retained due to the hydrophobic bases. The salt is not retained and is eluted first. The oligonucleotide may then be eluted using MeCN or MeOH-H_2O mixtures.

1.6.3 Oligonucleotide analysis

Nuclease digestion[49] is a commonly used technique for analysis of synthetic DNA. The oligonucleotide is incubated in a mixture of the enzymes snake venom phosphodiesterase and alkaline phosphatase providing a mixture of the monodeoxynucleoside constituents of the sequence. HPLC analysis allows the identity of each base in the sequence to be determined by coinjection with authentic samples of each nucleoside while peak area analysis allows for the base composition of the original sequence to be determined.

The actual code[50] of a purine sequence, for example, may be determined using specific chemical reactions to cleave a 5'-^{32}P labelled sequence at A and G bases respectively.[36, 51] When run in parallel, gel electrophoresis of both A and G specific reactions allows the sequence composition to be elucidated since longer segments of cleaved DNA move more slowly than short segments through the gel.

ESMS[52] is routinely used to rapidly identify sequences by providing accurate molecular weight determination of sequences without the need for radiolabelling and other less direct and time-consuming sample pretreatments such as nuclease digestion. For the purposes of this work, HPLC purification/analysis coupled with ESMS analysis were the techniques chosen to provide necessary unambiguous characterisation of purified oligonucleotides.

1.7 Drug attachment to TFOs: literature precedents

1.7.1 Where to attach the drugs to TFOs?

Since the mid 1980's, there have been many reports in the literature concerning the synthesis of modified oligonucleotides having a wide variety of ligands attached to them. Oligonucleotides have been functionalised and derivatised at the internucleotide linkages, the 3'-end, the carbohydrate component, the nucleobases and the 5'-end (Figure 1.7).[53] However, for this work, an important criterion based upon the mode of action of the drugs greatly reduced the number of sites which needed to be considered.

Following triplex formation of the TFO-drug conjugate with the duplex target, it was necessary for the drug to be in the duplex major-groove such that it could interact favourably with the proposed GGG target. Thus, it was considered that drug attachment to either a DNA base or the 5'-end of a sequence would position the drugs most favourably in the duplex major groove following triplexation.[54] Attachment at the 3'-end was also a possibility and this has been documented in the literature. For example,[55] cholesterol and 9-fluorenmethyl derivatives of adenosine nucleosides have been attached to CPG support allowing for synthesis of 3'-end modified sequences. The need for specific support

derivatisation using modified nucleosides in these cases made this approach less direct compared with attachment of **1** and **2** at the 5'-end and was thus not considered further in this thesis work.

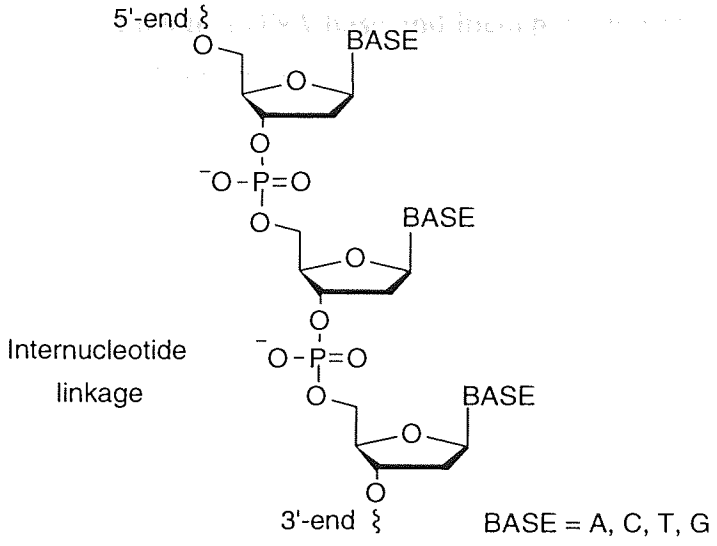


Figure 1.7 Potential sites for DNA modification; base, sugar, internucleotide linkage, 5'-end and 3'-end.

1.7.2 How to attach the drugs to TFOs?

Current methods[56] of oligonucleotide conjugate synthesis include utilisation of specific ligand phosphoramidites and modified CPG supports as well as post-synthetic modification of cleaved, deprotected oligonucleotides containing a reactive functionalised side chain. Occasionally, an activated nucleobase containing a good leaving group may be incorporated during DNA synthesis and displaced by an appropriately functionalised ligand. Post-synthetic conjugation is most commonly carried out in solution, however, there are a few reports concerning the modification of support-bound sequences.[43, 57, 58, 59, 60] This review will consider literature precedents for many of these routes after which, a comparison of the benefits and drawbacks of each will be given. Particular mention will be given to those procedures which have allowed for the incorporation of base-sensitive agents in oligonucleotides.

1.7.3 Direct incorporation of ligands: the phosphoramidite approach

Ligand incorporation using phosphoramidite methodology during DNA synthesis is the most direct method of modified-oligonucleotide synthesis. Generally, a phosphoramidite derivative of the ligand is synthesised which may only be incorporated at the 5'-end of the sequence or ligands are attached to a DNA base and so may be incorporated at any desired position within a sequence via phosphoramidite methodology. The examples which have

been included in this review focus on the methods used to incorporate a variety of ligands into oligonucleotides in order to demonstrate potential methods of drug-conjugate synthesis.

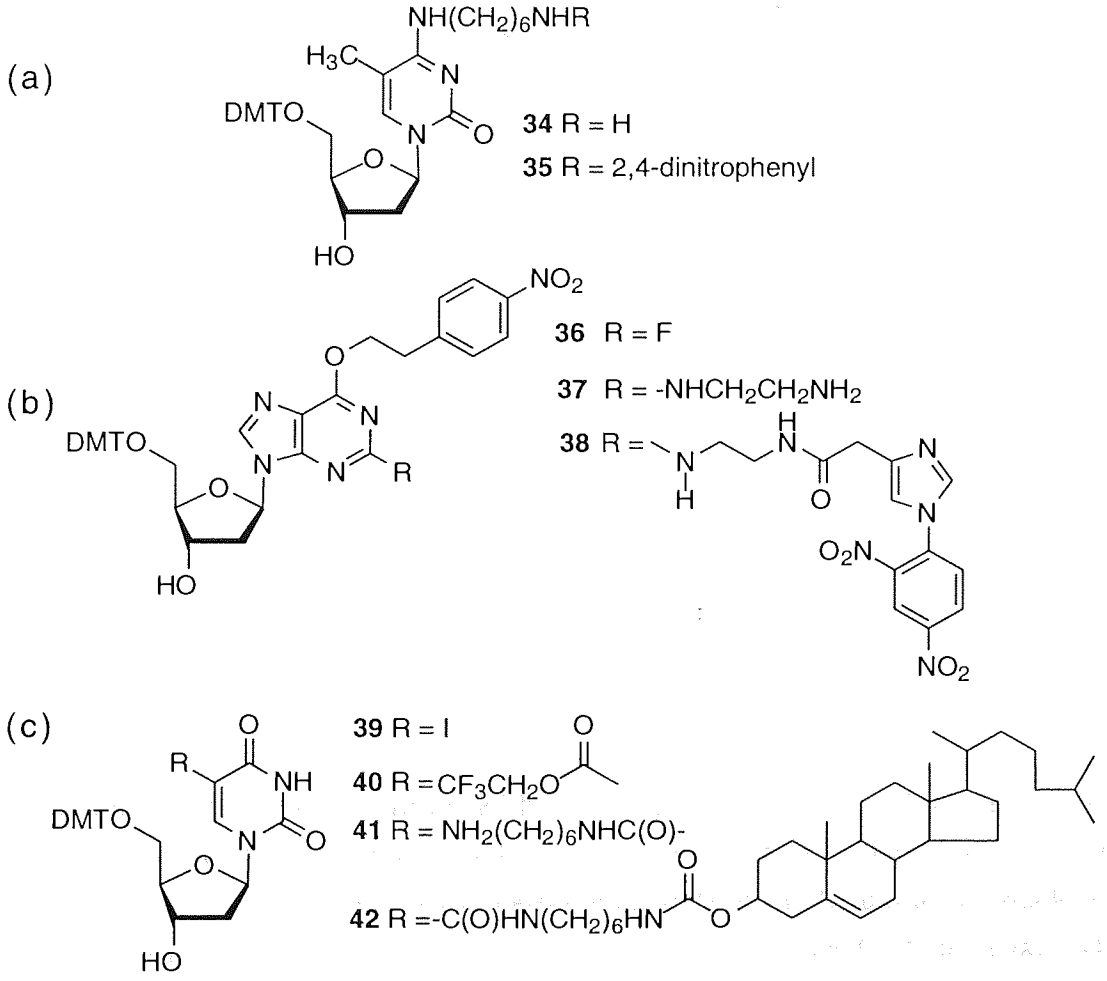
1.7.4 Attachment of ligands to a DNA base and incorporation via phosphoramidite methodology

There are few reports in the literature concerning the direct incorporation of ligand-derivatised DNA bases into DNA sequences. However some examples of this type do exist. Roget et al[61] converted 2'-deoxythymidine to its 4-thio derivative using phosphorus pentasulfide, which, following tritylation of the 5'-OH and displacement of the 4-thio group using an excess of 1,6-diaminohexane afforded **34** (Scheme 1.5 (a)). The free amino function of **34** was reacted with various activated markers including 1-fluoro-2,4-dinitrobenzene to afford **35**. Interestingly, in the case of the 2,4-dinitrophenyl derivative (**35**), ~30% of the 2,4-dinitrophenyl group was lost when **35** was treated to standard conc. NH_3 (aq) cleavage/deprotection conditions, following detritylation. Phosphoramidite derivatives of the modified nucleosides were used in automated DNA synthesis where high coupling efficiencies were observed (> 97%). Base-labile, PAC-protected phosphoramidites were recommended for use with the dinitrophenyl-labelled sequences, presumably to minimise exposure of the ligand to conc. NH_3 (aq).

Wang et al[62] displaced the F group of 2'-deoxyguanosine derivative **36** using ethylenediamine affording **37** (Scheme 1.5 (b)). The free amino function was derivatised with a 2,4-dintrophenyl protected imidazole using dicyclohexylcarbodiimide (DCC) coupling agent forming **38**. The corresponding phosphoramidite was then incorporated at internal positions of DNA sequences. Wang et al found that without the presence of the guanine O6 para-nitrophenethyl protecting group in **38**, the resulting phosphoramidite was too polar to be purified in good yield. Incorporation of the modified nuclosides was confirmed by enzymatic digestion analysis of the modified sequences.

Normura et al[63] attempted post synthetic, solution phase incorporation of cholesterol into a sequence containing an aminoalkyl tether attached to a base. However, the low solubility of the cholesterol in aqueous media prevented its attachment by this route. Instead, direct incorporation was attempted. Palladium-catalysed carbonylation of the iodouridine derivative **39** in the presence of carbon monoxide and CF_3CH_2OH afforded the trifluoroethoxycarbonyl compound **40** (Scheme 1.5 (c)). Reaction of **40** with 1,6-diaminohexane afforded the functionalised nucleoside (**41**). A carbonylimidazole derivative of cholesterol was reacted with **41** affording the desired nucleoside conjugate

(42). A phosphoramidite form of 42 was successfully incorporated into DNA sequences as evidenced by ESMS analysis and nuclease digestion analysis. This example illustrates the problems which may arise from differences in solubility between oligonucleotide and ligand.

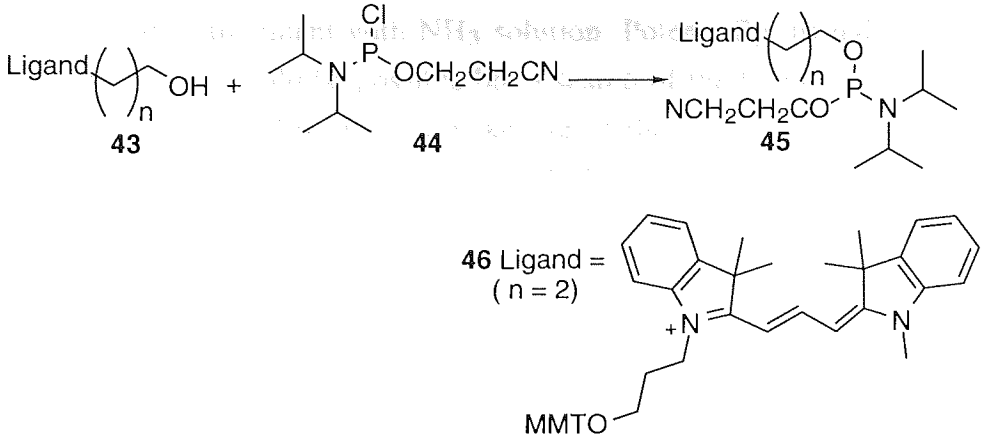


Scheme 1.5 Attachment of various ligands to purine and pyrimidine nucleosides.

1.7.5 Ligand attachment to the 5'-end using phosphoramidite methodology

This type of modified oligonucleotide synthesis is synthetically straight forward and so will not be dealt with in great detail. A hydroxy-alkyl derivative (**43**) of the ligand is reacted with the commercially available 2-cyanoethyl-*N,N*-diisopropyl chlorophosphoramidite (**44**) generating the ligand phosphoramidite (**45**) (Scheme 1.6) suitable for incorporation at the 5'-end of a sequence. The cyanine derivative[64] (**46**) is commercially available and is of interest here because, once incorporated into a sequence, it is unstable towards prolonged treatment with conc. NH_3 (aq) at elevated temperatures. The user is therefore advised to

limit exposure of the ligand to this reagent (2 h, 55 °C) by using monomers containing base-labile *N*-protecting groups.



Scheme 1.6 General synthesis of ligand phosphoramidites.

1.7.6 Post-synthetic conjugate formation

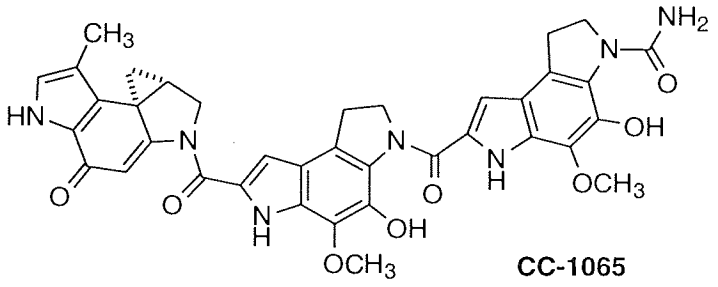
Post-synthetic conjugation of any auxiliary to DNA may either be carried out in solution or while the oligonucleotide is still support-bound. Of the two approaches, solution phase sequence derivatisation methods are by far the more popular, however, both approaches are of interest here. Solution phase methods involve the direct insertion of a reactive, nucleophilic function (for example, amino), on an alkyl tether, into a sequence using phosphoramidite methodology, followed by cleavage/deprotection and purification of the modified sequence. Any suitable ligand may then be attached to the oligonucleotide by coupling a suitable derivative (usually carboxylic acid) of the ligand to the nucleophilic function on the oligonucleotide using a coupling agent (DCC, benzotriazolyloxy-tris[pyrrolidino]-phosphonium hexafluorophosphate (PyBOP)). Subsequent conjugate purification usually takes place in two steps. Size-exclusion chromatography is often used initially to separate the conjugate from excess ligand and from other unreacted materials (coupling agent). Preparative HPLC is then used to separate the conjugate from unlabelled sequences.

1.7.7 Solution phase formation of conjugates

Postsynthetic 5'-end derivatisation of oligonucleotides is a relatively easy, standardised procedure compared with ligand conjugation to oligonucleotides at internal positions due to the commercial availability of phosphoramidites such as **47**[65] and **48**[66] (usually, n = 6, Scheme 1.7 (a)). These phosphoramidites may be coupled to the 5'-end of a sequence as efficiently as standard nucleoside phosphoramidites[67] which, following cleavage and

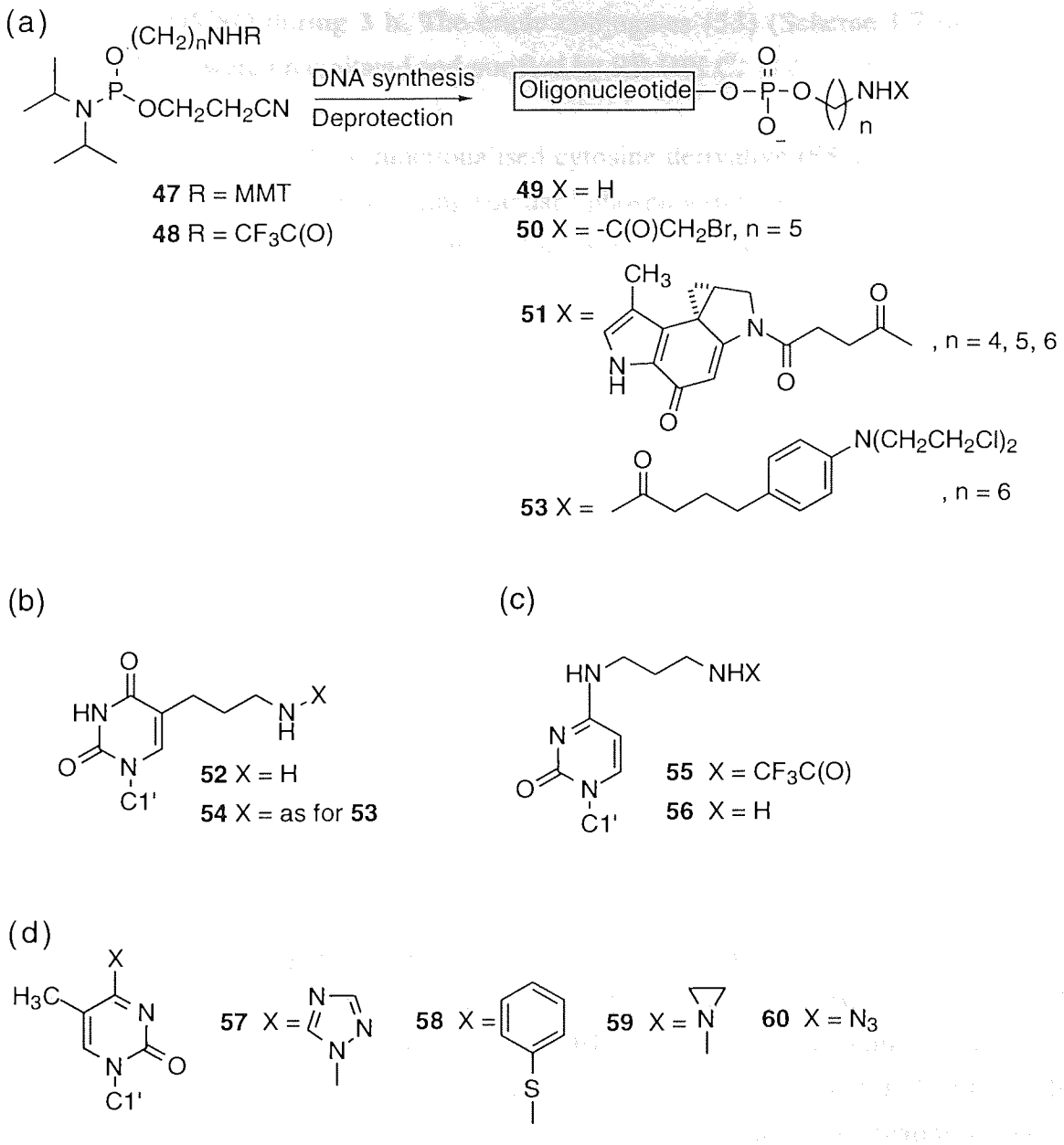
deprotection, affords the functionalised sequence (**49**) (Scheme 1.7 (a)). In the case of **48**, the trifluoroacetyl group is cleaved using standard NH_3 (aq) deprotection conditions (conc. NH_3 (aq), 20 h at 60 °C).[66] However, the MMT group of **48** is acid labile and thus may be cleaved before or after treatment with NH_3 solution. Potentially, ligand coupling to the support-bound sequence would be possible but a search of the literature has not revealed any examples of this type. However, to make use of the lipophilic nature of the MMT group, it is more commonly cleaved (80% AcOH, 1 h, r.t.) after standard cleavage, deprotection and reversed-phase HPLC purification of the sequence.

Grant et al[67] incorporated the phosphoramidite **48** (n = 5) at the 5'-end of a sequence which, following standard cleavage/deprotection, was purified by two rounds of preparative gel electrophoresis. The modified sequence was reacted with N-hydroxysuccinimidyl bromoacetate (250 eq) during 5 min at pH 8.5 to afford the bromoacetyl alkylating conjugate **50**.



CC-1065

Lukhtanov et al[68] synthesised phosphoramidite **47** (MMT n = 4, 5 or 6) according to Connolly et al[65] to incorporate a variety of amino-linker arms at the 5'-end of sequences (Scheme 1.7 (a)). The tetrafluorophenoxy ester derivative (5-20 eq) of the cyclopropapyrrolonindole (CPI) alkylating subunit of the antitumor antibiotic CC-1065 was coupled during 16 h to the triethylammonium (TEA) salts of the purified amino-linker oligonucleotides, under anhydrous conditions in DMSO, to give the desired conjugate **51**. In total, there were four purification steps which included HPLC purification of the MMT-ON linker sequence followed by purification of the detritylated sequence. Following conjugation, the crude product was precipitated from the reaction mixture and then HPLC purified. Nuclease digestion was used to confirm incorporation of the alkylating agent.



Scheme 1.7 Post synthetic modification of DNA using: (a) linker phosphoramidites at the 5'-end, (b)-(c) linker nucleosides and (d) activated nucleosides.

Belusov et al[54] incorporated chlorambucil both internally and at the 5'-end of a sequence. For the 5'-end modification, **47** (MMT, n = 6) (Scheme 1.7 (a)) was incorporated at the 5'-end during synthesis while the uridine derivative (**52**) (Scheme 1.7 (b)) was incorporated to allow internal modification. For the purposes of solubility, the fully cleaved/deprotected oligonucleotides were converted to either TEA or tributylammonium (TBA) salts using HPLC prior to reaction of the free amine with the tetrafluorophenoxy ester of chlorambucil

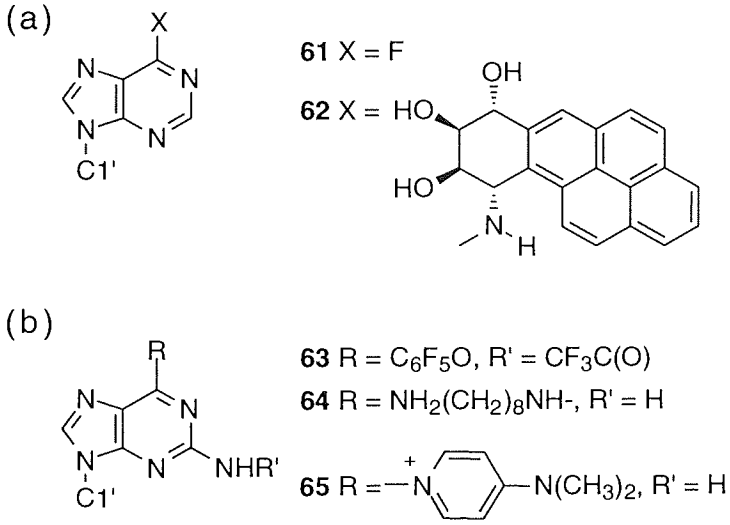
in anhydrous DMSO during 3 h. The crude conjugates (**53**) (Scheme 1.7 (a)) and (**54**) (Scheme 1.7 (b)) were precipitated and purified by RP-HPLC.

Telser et al[69] incorporated the functionalised cytosine derivative (**55**) (Scheme 1.7 (c)) at an internal position in sequences using standard phosphoramidite methodology in order to attach molecular labels such as biotin, fluorescein and pyrene. Compound (**55**) was synthesised via a bisulfite-catalysed transamination reaction involving 1,3-diaminopropane and cytidine afterwhich the free amino group was trifluoroacetyl protected. An excess (~150 eq) of the activated (N-hydroxy succinimide ester, sulfonyl chloride or isothiocyanate) label in DMF was reacted with the deprotected, purified oligonucleotide containing **56** dissolved in sodium borate buffer (pH 9.3) during 12 h. Incidentally, all labels were found to be rather insoluble in the buffer used. The conjugates were purified by size-exclusion chromatography followed by liquid chromatography. Ligand incorporation was also evident by UV analysis of the conjugates. It was suggested that such coupling reactions should be performed at ~ pH 9 to ensure that the primary amino group was completely deprotonated provding a better nucleophile. It was observed that reaction of activated labels occurred only with the functionalised oligonucleotides and not with the unmodified sequences. It was suggested that the primary amine was a better nucleophile than the aryl amines on the DNA bases as the latter can exist as imino tautomers. Also, the linker arm made the primary amino function more accessible.

In order to synthesise oligonucleotides containing base-sensitive groups, Zheng et al[70] used phosphoramidite methodology to incorporate the thymidine derivative (**57**) (Scheme 1.7 (d)) internally in a sequence synthesised using BPA^t-N-protected monomers. The triazolo group was converted to the phenylthio derivative (**58**) by treating with $PhS^-[EtNPr^i_2]H^+$ overnight. Cleavage and deprotection of the sequence was achieved by treating with conc. NH_3 (aq) at r.t. during 2 h. Longer exposures would have resulted in displacement of the phenylthio group by NH_3. The base sensitive derivatives (**59**) and (**60**) were synthesised by treating the purified oligonucleotide with aqueous aziridine (r.t., 20 min) and heating with NaN_3 in 50% DMF/H_2O (35 °C, overnight), respectively. Ligand incorporation prior to treatment with NH_3 resulted in degradation of the conjugates producing sequences containing 5^mC or thymine which formed by ammonolysis and hydrolysis of the phenylthio group respectively. It was claimed that this method provided a way of incorporating extremely reactive ligands into sequences since no post synthetic treatments were required after substitution of the phenylthio group by the ligand.

1.7.8 Post synthetic modification of solid-supported oligonucleotides

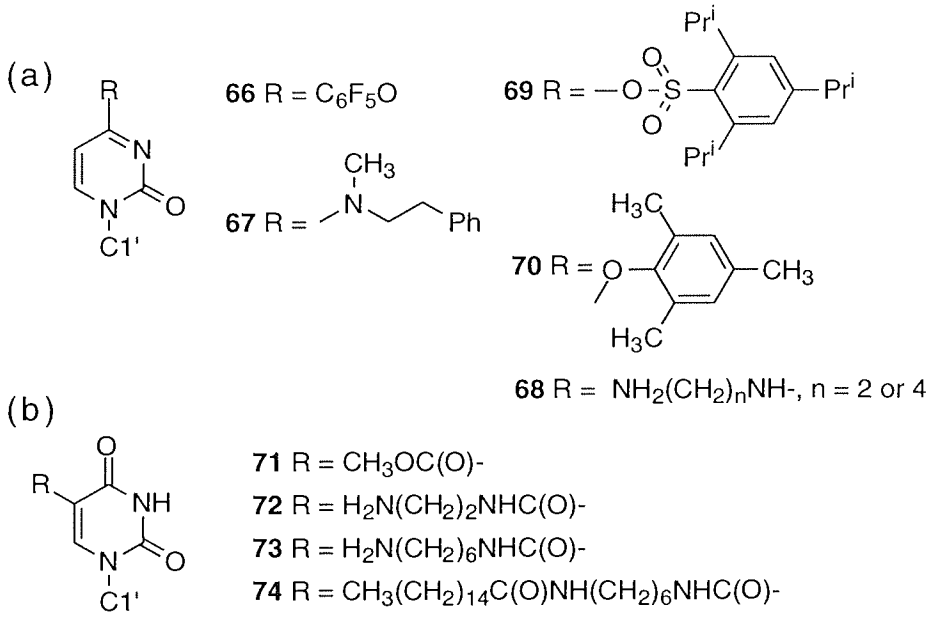
Kim et al[57] synthesised polycyclic aromatic hydrocarbon (PAH) adducts of oligonucleotides using the adenosine derivative (**61**) (Scheme 1.8 (a)) which was synthesised from 6-chloropurine by reaction with KF in the presence of Me_3N. Following phosphoramidite incorporation of **61** during DNA synthesis, PAH derivatives such as **62** were synthesised by heating the fully protected CPG-bound sequence in the presence of solutions of the PAH's in DMF and Et_3N (53 °C during 5 days for **62**). It was reported that excess of the relatively expensive PAH starting materials were recovered by washing with MeOH. The conjugate sequences were cleaved and deprotected using conc. NH_3 (aq) (60 °C, 8 h).



Scheme 1.8 Modified nucleosides (a), (b) which have been incorporated in to oligonucleotides and their derivatives resulting from subsequent reaction of the support-bound sequence with various nucleophiles.

Gao et al[58] used phosphoramidite methodology to incorporate the activated guanosine derivative **63** (Scheme 1.8 (b)) into oligonucleotides at both internal and the 5'-terminal positions. **63** was synthesied by reaction of 2'-deoxyguanosine with trifluoroaceticanhydride in pyridine followed by treatment with pentafluorophenol. It was possible to displace the pentafluorophenoxy groups from the internal and 5'-terminal modified bases of the fully protected support-bound sequences using 1,8-diaminoocatne (10% solution of 1,8-diaminooctane in MeCN at 60 °C during 39 h) and MeCN saturated with DMAP (60 °C during 2 days). Cleavage and deprotection under mild conditions was achieved using conc. NH_3 (aq) (r.t. during 3 days), affording sequences containing the derivatised bases (**64**) and (**65**).

Wallis et al[59] synthesised the C4-activated uridine derivative **66** (Scheme 1.9 (a)) similar to Gao et al[58] and incorporated it internally into a sequence using phosphoramidite methodology. It was found that the pentafluorophenoxy group could be displaced, although inefficiently, from the fully protected support-bound oligonucleotide using *N*-methyl-*N*-phenethylamine at 60 °C during 20 h. Standard treatment with conc. NH_3 (aq) afforded the fully cleaved and deprotected oligonucleotide sequence containing the cytosine derivative (**67**).



Scheme 1.9 Modified nucleosides (a), (b) which have been incorporated in to oligonucleotides and their derivatives resulting from subsequent reaction of the sequence with various nucleophiles.

MacMillan et al[43] incorporated the uridine derivative (**70**) (Scheme 1.9 (a)) at an internal position within an oligonucleotide. Compound (**70**) was synthesised via a route which involved activation of the uridine C4 carbonyl by conversion to the *O*-(2,4,6-triisopropylbenzenesulfonyl) analogue (**69**) which was subsequently elaborated to **70** using trimethylphenol. The use of phosphoramidite monomers dC^{Bu^i}, dA^{PAC} and dG^{PAC} allowed full cleavage and deprotection of the sequence using conc. NH_3 (aq) (r.t., 7 h) without cleavage of the trimethylphenoxy group of **70**. The trimethylphenoxy group was then displaced by heating the sequence seperately with amines in MeCN (1-22 M) at 65 °C during 14 h providing the modified bases such as **68** within the sequence.

Haginoya et al[60] incorporated the C5-activated uridine derivative (**71**) (Scheme 1.9 (b)) into oligonucleotides at various positions. Treatment of the fully protected support-bound

oligonucleotides with methanolic solutions of ethylenediamine or 1,6-hexanediamine (45 °C, 12-14 h) afforded the appropriate (*N*-aminoalkyl)-carbamoyl derivatives (**72**) and (**73**) as well as effecting full deprotection and cleavage of the oligonucleotides from the support. Normura et al[63] synthesised sequences containing (**73**) by a similar route and derivatised the amino function with various ligands. For example, a solution of *N*-hydroxysuccinimidyl palmitate in DMF was added to an equal volume of the oligonucleotide dissolved in basic aqueous solution (pH 10.3) during 12 h at r.t. The purified palmitic acid derivative (**74**) was isolated following HPLC and size exclusion chromatograpy and analysed by ESMS and nuclease digestion analysis. Haginoya et al[60] maintained that this route towards linker arm incorporation had the following advantages over other methds:

1. Protection and deprotection of the aminoalkyl nitrogen was unnecessary.

2. The route avoided the solubility problems[62] encountered during the synthesis of aminoalkyl derivatised nucleoside phosphoramidites.

3. In instances where the optimal linker arm length is unknown, several linkers of various lengths could be incorporated using this method without the need for the separate construction of the appropriate nucleoside phosphoramidite units and their time-consuming, independent incorporation into oligonucleotides.

Obviously, in order to incorporate an auxiliary, solution phase coupling would be necessary unless an amino alkyl tether could be attached to the ligand, in which case solid-supported coupling of the ligand to the oligonucleotide could be possible.

1.7.9 Comparison of methods for oligonucleotide-conjugate formation

Direct incorporation of a ligand at the 5'-end via phosphoramidite methodology has the following advantages:

1. There are no solubility problems associated with the often hydrophobic ligand and hydrophilic DNA sequences which are commonly encountered with post synthetic coupling.

2. Protection of amino-alkyl tethers is unnecessary.

3. There is no problem of over-reaction where ligand incorporation occurs at undesired sites within the sequence.

4. Fewer chromatographic purification steps are required:

   (a) no need for purification of amino-linker sequenece (**49**) (Scheme 1.7)
   (b) no need for size exclusion chromatography since excess ligand is washed away during synthesis.

The direct phosphoramidite approach is desirable since small amounts of the ligand phosphoramidite are used and coupling efficiencies in excess of 98% may be achieved. However, the phosphoramidite approach can be time-consuming and laborious since multiple specific ligand-phosphoramidite synthesis is necessary, each of which has to be seperately incorporated into individual DNA sequences. Also, synthesis of a ligand phosphoramidite may be hindered due to the highly polar nature of some phosphoramidites, making elaborate synthesis necessary. However, it was felt that the advantages of this approach far outweighed the disadvantages and was thus attempted in this work. Where the drugs (**1**) and (**2**) are concerned, conjugation both at the 5'-end and within the sequence via the phosphoramidite approach necessitated alternative cleavage and deprotection conditions since the drugs are not stable towards standard conc. NH_3 (aq) treatment. Section 1.10 comprises a brief but comprehensive review of potentially useful novel supports.

Post synthetic modification of oligonucleotides has one main advantage over the phosphoramidite approach in that several derivatives may be generated from the synthesis of a single, suitably functionalised sequence. Also, the elaborate synthesis of individual substrate phosphoramidites is unnecessary. Although popular, the solution phase approach towards post-synthetic derivatisation of oligonucleotides suffers a disadvantage in that, overall, four chromatographic purification steps are usually necessary making conjugate synthesis a time-conuming process and potentially poor yielding. The solid phase approach however, by its very nature, could considerably reduce the time of the process as well as increase the overall yield of conjugate for the following reasons:

1. Efficient use is made of the modified DNA sequence, that is to say, the sequence is both synthesised and conjugated to the substrate while still support-bound.

2. Conjugate purification is simplified as excess coupling agent/substrate can be removed with ease by washing the support.

3. Ease of handling of the CPG-bound sequence.

4. Potentially higher yielding as only a single HPLC purification of the cleaved conjugate is necessary, thus minimising transference losses. As conjugates are often more hydrophobic than unconjugated/failed sequences, RP-HPLC purification is easier.

For these reasons it was considered that post-synthetic, solid-supported oligonucleotide-drug conjugate synthesis was the method of choice for both 5'-end and internal attachment of the drugs.

1.8 Determination of a DNA alkylation event

The site of alkylation within a purine sequence may be determined by heating the 5'-^{32}P labelled sequence containing the alkylated lesion in piperidine (55 °C, 12 h) which causes depurination of the alkylated base, ultimately leading to strand scission at that point.[38] The mobility of the labelled fragment in gel ectrophoresis allows the point of alkylation to be determined when compared with A- and G-specific chemical sequencing reactions of the non-alkylated sequence.

A popular method used for determination of a cross-linking event is the gel retardation assay which is run under denaturing conditions where duplexes and triplexes cannot exist. Covalently bound (cross-linked) sequences, of which one is ^{32}P radiolabelled at the 5'-end, have a lower gel mobility compared with the single radiolabelled strand due to their larger size.

Although techniques such as the gel retardation assay suggest that modification of a duplex target has taken place, it does not directly allude to the identity of the reaction product.[52, 71] When multiple products result from a cross-linking reaction, resolution and identification of the components would be difficult using a technique such as gel retardation. Mass spectrometric techniques have been comparatively little used to identify the products of cross-linking/alkylation reactions resulting from the binding of TFO-alkylator conjugates to target duplex sites.

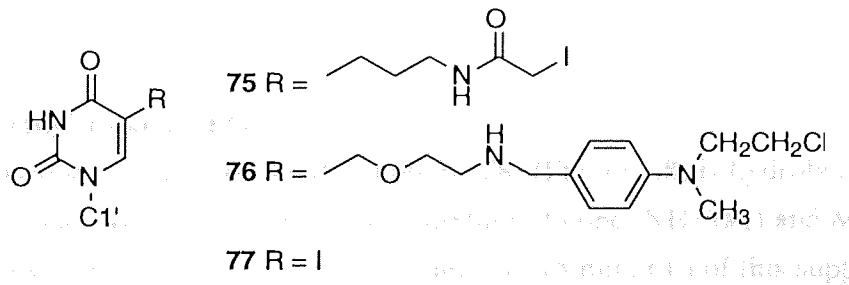
1.9 DNA alkylation using TFO-conjugates of DNA alkylating agents

Shaw et al[72] used the *N*4, *N*4-ethano 5^{m}C derivative (**59**) (Scheme 1.7) in a 5^{m}C- and T-containing TFO to target a polypurine segment of a 5'-end ^{32}P labelled sequence of a duplex. Evidence of a cross-linked species was observed using denaturing gel electrophoresis. When the target duplex was incubated (pH 7.2, 37 °C, 16 h) seperately

with the unmodified and modified TFOs, a slow migrating species was only observed in the gel for the latter when compared with the labelled single stranded target sequence. Treatment of cross-linked sequences with pyrrolidine resulted in depurination of N7 alkylated Gs and chain cleavage at the depurinated site. Comparison of the gels of a G-specific dimethyl sulphate sequencing reaction and that of the pyrrolidine incubation indicated the location of the alkylation site in the polypurine target.

Grant et al[67] incubated (24 h, 37 °C) a 15-base TFO derivatised at the 5'-end with the bromoacetyl group, with a 645 base-pair duplex containing 3 consecutive guanine bases adjacent to the 3'-end of the TFO binding site. The yield of single strand cleavage after piperidine treatment and gel electrophoresis was 80% at the G base directly adjacent to the binding site and 7% and 1% at the G bases one and two bases removed illustrating the specificity of alkylation.

Meyer et al[73] showed that when the uridine derivative (**75**), containing an iodo acetyl alkylating group, was incorporated into a mixed A, C, T and G containing duplex forming sequence, site specific alkylation at the N7 of a G base two bases removed from the modified nucleoside in the target sequence was achieved following incubation at 37 °C, as evidenced by the piperidine method described above. Alkylation of the target was not as pronounced when incubation was carried out at temperatures above the T_m of the duplex, indicating that duplex formation was necessary for alkylation to occur. Further evidence in support of this came from the fact that cross-linking was inhibited in the presence of an excess of the duplex forming sequence lacking the electrophilic side-arm.



Scheme 1.10 Nucleosides which have been incorporated into oligonucleotides and successfully used to effect site-specific alkylation of target sequences.

Khalimskaya et al[74], incorporated the uridine derivative (**76**) (Scheme 1.10), containing the alkylating 4-(*N*-2-chloroethyl-*N*-methylamino)benzyl group, at the middle of a 7-mer duplex-forming sequence. Incubation of stoichiometric quantities of the conjugate

containing **76** with a 16-mer target at 22 °C during 48 h resulted in 83% modification of the target with predominant alkylation occurring at a single G residue located adjacent to the modified base on the target strand. The extent and points of modification were determined by gel electrophoresis after cleavage of the target at the alkylated purine residue (10% aqueous piperidine, 100 °C, 30 min) and comparison with the gel of a G/A sequencing reaction of the target.

Recently, Wong et al[75] photochemically cross-linked an oligonucleotide containing 5-iodouridine (**77**) (Scheme 1.10) to a protein (M.EcoRI) on nM and μM scales. The cross-linked products were seperated by gel electrophoresis and desalted. Analysis of the cross-linked conjugate (molecular weight (MW) ~ 42000) by ESMS confirmed its structure with a difference of 1.4 mass units between the calculated and measured masses.

For the purposes of this project, ESMS was considered a useful technique for providing rapid, accurate evidence of alkylation/cross-linking of duplex DNA targets. Obviously, a technique such as the piperidine cleavege method would thereafter be necessary to locate the site of alkylation.

1.10 Novel solid supports

A new set of cleavage conditions, compatible with the base-sensitive oligonucleotide conjugates of drugs (**1**) and (**2**) was necessary to allow incorporation of the drugs into sequence via phosphoramidite methodology. The following is a brief review of the published procedures and CPG supports which have been developed to achieve cleavage of sequences from solid supports under milder conditions than the standard conc. NH_3 (aq) method.

1.10.1 Succinyl-linked CPG

Commonly used succinate-linked CPG support (**78**) (Figure 1.8) is hydrolysed slowly[76, 77] (~ 1 h) using conc. NH_3 (aq). AMA[44] (1:1 mixture of conc. NH_3 (aq) and $MeNH_2$ (aq)) is a commercially available solution for rapid cleavage (5 min, r.t.) of this support. However, the base-labile drugs (**1**) and (**2**) are not stable under either AMA or NH_3 conditions and so a method was sought for cleaving a sequence from a CPG support under mild, non-basic drug-compatible conditions.

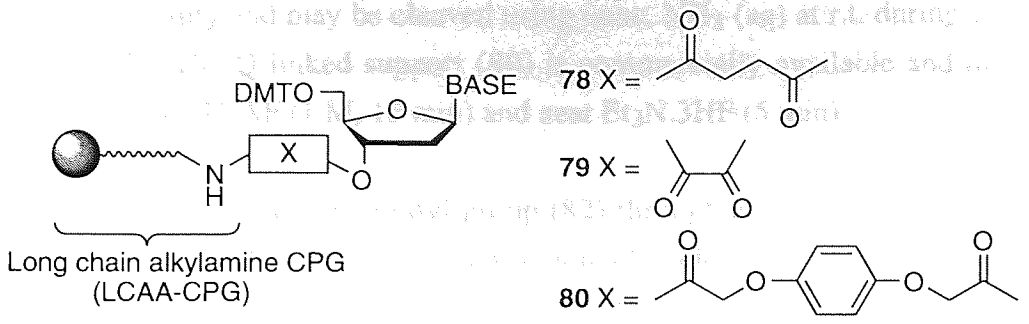
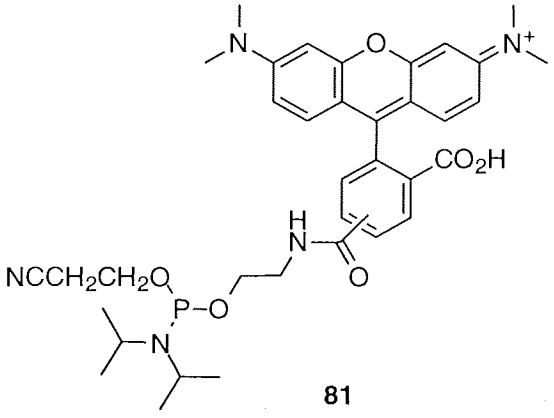


Figure 1.8 Base-labile CPG supports.

During recent years, a number of different routes has been developed which provides more rapid and mild cleavage of the succinate-linked support (**78**) (Figure 1.8).



81

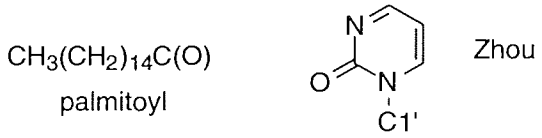
Rhodamine dye is usually incorporated into DNA sequences following cleavage and deprotecion of the sequence since exposure of the dye to NH_3 (aq) causes degradation. Woo et al[78] argued that more efficient attachment of the dye could be achieved using the dye-phosphoramidite (**81**). Following incorporation of the dye at the 5'-end using phosphoramidite methodology, a 1:1:2 mixture (by volume) of Bu^tNH_2-MeOH-H_2O was used to cleave the oligonucleotide-dye conjugate from the support during 1 to 2 h at r.t. The sequence was fully deprotected by heating the cleavage solution at 80 °C during 20 to 60 min without degradation of the attached dye.

1.10.2 The Oxalyl and Q-linked CPG supports

Routes towards mild, rapid cleavage of sequences from CPG supports have involved the use of more labile alternatives to the commonly used succinate support (**78**). The oxalyl-linked[79] and hydroquinone (Q[77, 80])-linked supports (**79**) and (**80**) (Figure 1.8) have been developed in recent years for the purpose of synthesising oligonucleotides containing base-

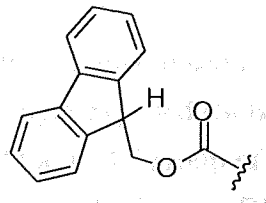
sensitive functionality and may be cleaved using conc. NH_3 (aq) at r.t. during 1 and 2 to 5 min respectively. The Q-linked support (**80**) is commercially available and may also be cleaved using either TBAF (1 M, 15 min) and neat $Et_3N.3HF$ (5 min).

Polushin et al[81] attached the palmitoyl group (**82**) through a base-sensitive ester bond to the terminal 5'-hydroxyl of fully protected oligonucleotides bound to oxalyl-linked CPG. The conjugate sequences containing base-labile BPA^t protecting groups on A, C and G, were treated with neat ethanolamine during 10 min. which afforded the completely cleaved/deprotected sequences. However, although treatment with conc. NH_3 (aq) effected cleavage of the sequence from the support during 2 min, treatment during 2 h at r.t. was necessary to fully deprotect the sequence. This treatment was found to almost completely hydrolyse the palmitoyl ester bond, a problem solved by the use of the oxalyl support (**79**).



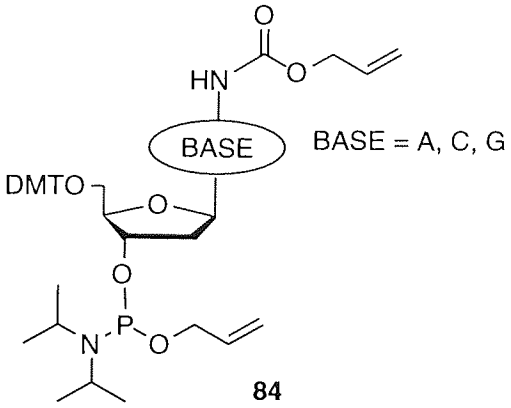
82 **83**

Incorporation of the fluorescent pyrimidine analogue 2-pyrimidinone (**83**) into oligonucleotides has been problematic since degradation (~ 30%) occurs readily during treatment with conc. NH_3 (aq) (1 h, r.t.) through attack of NH_3 at the C5-C6 double bond.[82] Zhou et al[83] synthesised sequences comprising only **83** on oxalyl support (**79**) using 5% NH_3 (aq) in MeOH (70 min, r.t.) for cleavage and deprotection of the oligonucleotide. It was found that degradation of the 2-pyrimidinone base had not occurred. It was suggested that fluorenylmethoxycarbonyl (Fmoc)-protected nucleoside monomers of A, C and G could be used to incorporate multiple residues of **83** into sequences containing natural bases using oxalyl support (**79**) since the Fmoc group may be cleaved using conditions[84, 85] under which the 2-pyrimidinone base is stable such as 1:1 DIPA-MeOH at 55 °C during 15 h or 10% anhydrous piperidine-MeOH at r.t. during 36 h.



Fmoc

Alul et used the oxalyl support to synthesise oligonucleotides with base protecting groups intact. The fully protected support-bound sequence was decyanoethylated by treating with a hindered base such as diisopropylamine (DIPA) during 14 h. The decyanoethylated sequence was cleaved from the support using 20% NH_3 (aq) in MeOH (2 min) affording the fully base-protected (A^{Bz}, C^{Bz}, G^{Bu^i}) sequence.



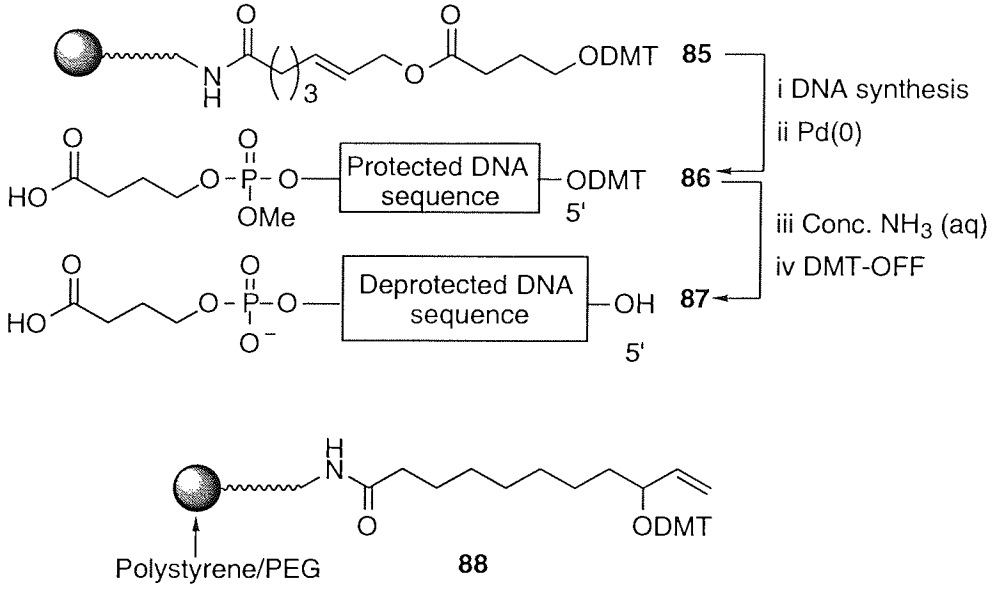
84

Hayakawa et al[86] devised a method of synthesising fully deprotected, support-bound oligonucleotides. They used phosphoramidite monomers of type (**84**) during DNA synthesis on succinate support in which allyloxycarbonyl (AOC) and allyl serve as base and phosphorus protecting groups respectively. Complete deprotection of the sequence was achieved by treatment of the support-bound sequence with a mixture of triphenylphosphine (PPh_3), tris(dibenzylacetone)dipalladium(0)-chloroform ($Pd_2(dba)_3$-$CHCl_3$) and n-butylammonium formate (Bu^nNH_2-HCO_2H) in THF at 50 °C during 0.5 to 1 h. The sequence was then cleaved from the support by treating with conc. NH_3 (aq) at r.t. during 2 h. It was claimed that isolation of target DNA sequences was simplified using this method since all unbound contaminants and cleaved protecting groups, including unreacted reagents, could be removed by simple washing of the support. Since this work, solid supports containing allyl linkages have been synthesised which allow for very mild palladium-catalysed cleavage.

1.10.3 Allyl-linked supports

Matray et al[87] synthesised fully protected, mixed sequences which, when cleaved from the allyl-linked CPG (**85**), contained a single exposed (Scheme 1.11) carboxylic acid function at the 3'-end. Following cleavage of the DMT group of **85**, the first OMe-phosphoramidite was added to the exposed hydroxyl group. Following DNA synthesis, mild Pd(0)-catalysed cleavage of the sequence from the support was achieved using bis(diphenylphosphino)ethane and $Pd_2(dba)_3$-$CHCl_3$ in Bu^nNH_2-HCO_2H (1.2 M) at 55 °C during 7 h which afforded the fully protected sequence (**86**) containing a 3'-alkyl

carboxylic acid. It was suggested that ligand conjugation to such modified, protected sequences would prevent the occurrence of unwanted reactions at the exocyclic amino functions of the bases or the phosphate groups. Also, the improved solubility of protected oligonucleotides in organic solvents would aid solution-phase coupling. Treatment of **86** with conc. NH_3 (aq) (10 h, 55 °C) followed by removal of the 5'-DMT group afforded the fully deprotected sequence (**87**).



Scheme 1.11 Allyl-linked solid supports which may be cleaved under mild conditions using Pd(0)-mediated catalysis.

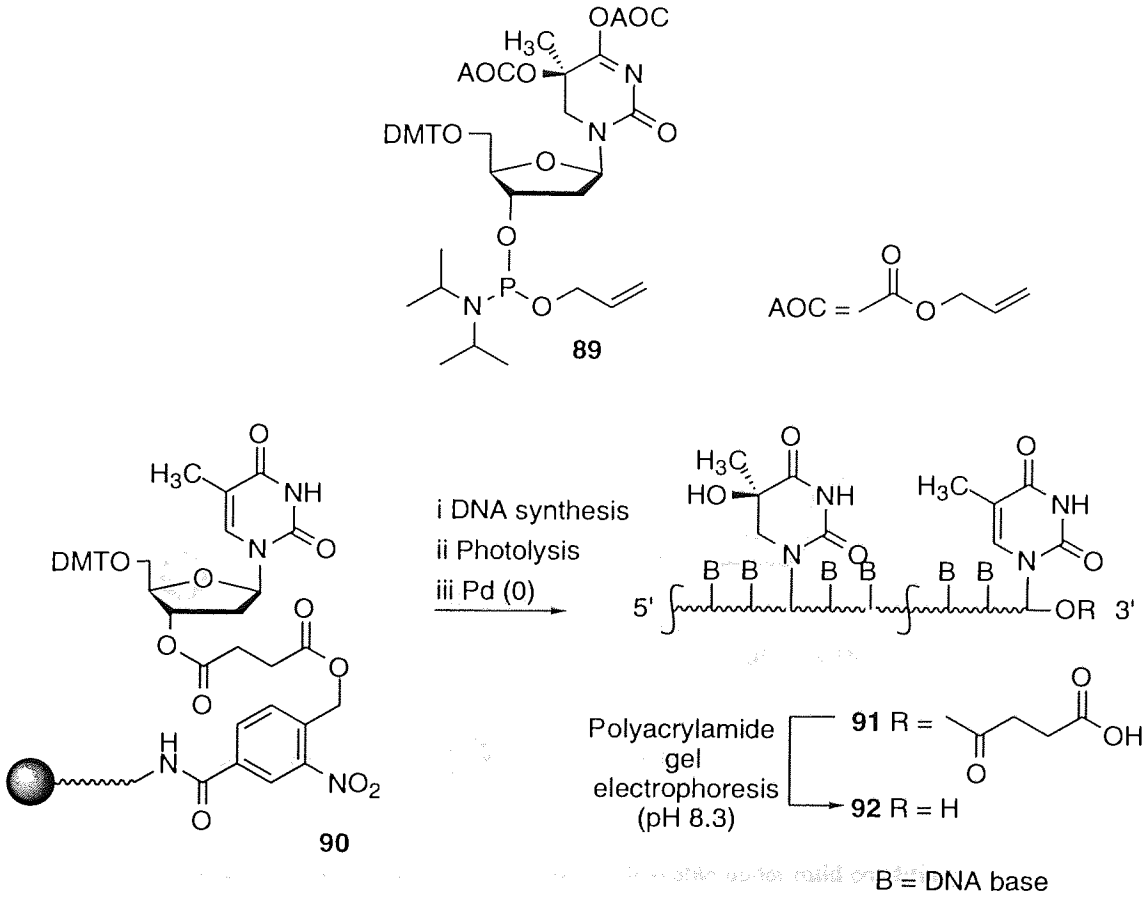
Similarly, Zhang et al[88] used the allyl-linked polystyrene/PEG supports (**88**) (Scheme 1.11) to synthesise RNA sequences. As for **85**, monomers were added to the free hydroxyl which was generated by cleavage of the DMT group of **88**. The support was found to be stable towards a 3:1 mixture of conc. NH_3 (aq) in MeOH at r.t. during 16 h and to $Et_3N.3HF$ at 58 °C during 4 h with the result that the sequences could be fully deprotected while still support-bound. Release of the oligoribonucleotide from the support was achieved using tetrakis(triphenylphosphine)palladium and Bu^nNH_2-HCO_2H at 60 °C during 2 h. It was suggested this support would be suitable for the synthesis of base-sensitive nucleic acids.

1.10.4 Photolabile supports

Matray et al[89] incorporated the base sensitive nuceolside (*5R*)-5,6-dihydro-5-hydroxythymidine directly into an oligonucleotide using its phosphoramidite form (**89**) and the DMTdT-derivatised photolabile support (**90**) (Scheme 1.12). The DNA bases A, C, T

and G were incorporated using monomers of type (**84**) where AOC and allyl served as base and phosphorus protecting groups respectively. This allowed for complete support-bound deprotection of the sequence using PPh_3 and $Pd_2(dba)_3$-$CHCl_3$ in Bu^nNH_2-HCO_2H at 55 °C during 1 h. The oligonucleotide was cleaved from the support using photolysis (λ_{max} = 400 nm) during 3 h affording the 3'-modified sequence (**91**). During polyacrylamide gel electrophoresis (PAGE) purification, the 3'-*O*-succinate moiety was cleaved under the mildly basic conditions (pH 8.3) affording the fully deprotected sequence (**92**).

A drawback of the photolabile supports[87] was realised when a decrease in yield of oligonucleotide was observed as the length was increased from 20 to 40 nucleotide residues. The reduced yield was attributed to reduced support-cleavage due to increased competition of the oligonucelotide for light during photolysis compared with the *o*-nitrobenzyl chromophore of **90**.

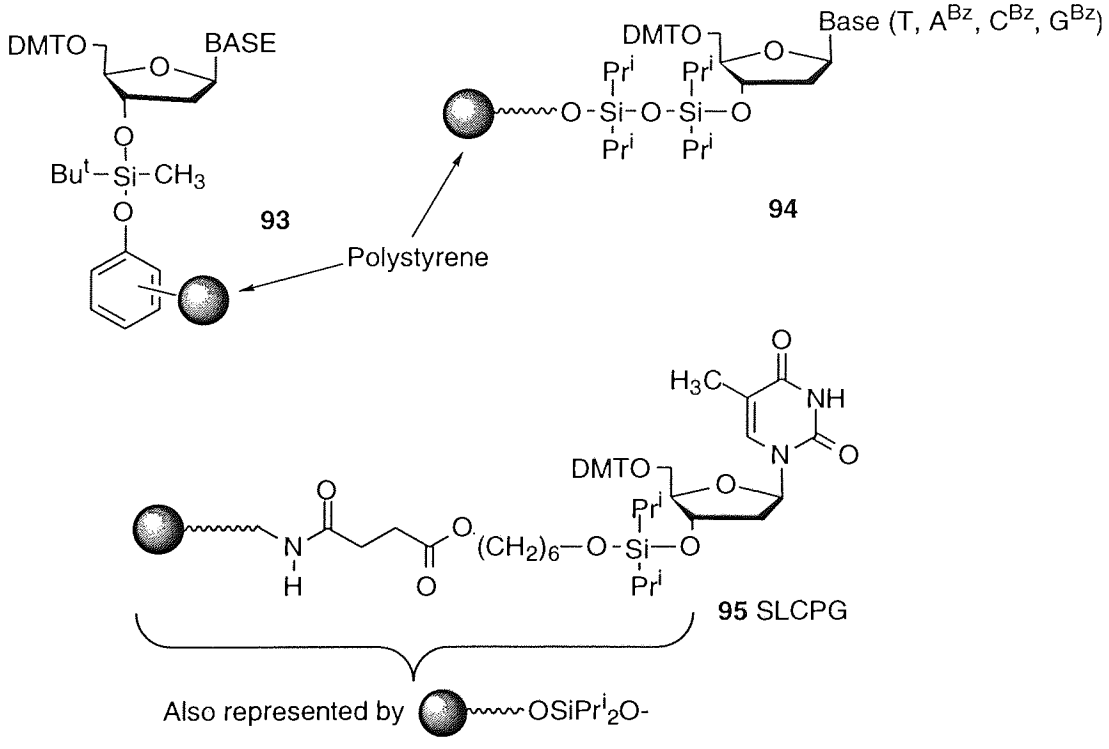


Scheme 1.12 The photolabile support (**90**) and its use in the incorporation of the base-sensitive nucleoside (5*R*)-5,6-dihydro-5-hydroxythymidine into an oligonucleotide using the phosphoramidite (**89**).

1.10.5 Silyl-linked supports

Holmberg[90] used a variety of silyl-linked polystyrene supports for DNA synthesis, of which **93** (Scheme 1.13) is an example. Sequences synthesised using phenoxyacetyl *N*-protected monomers were fully deprotected using conc. NH_3 (aq) (70 °C during 60 min) while still bound to the support. Cleavage of sequences from the support was achieved using either 0.1M TBAF in DMF at 70 °C during 30 min or 1.0 M NaOH at 70 °C during 15 min.

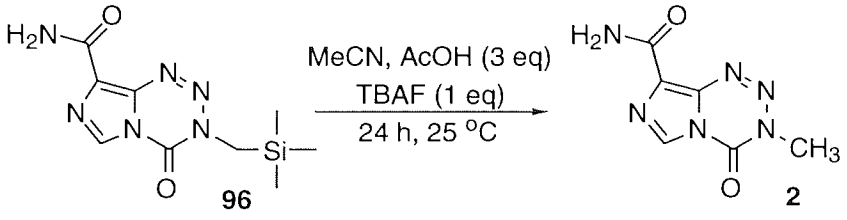
Similarly, Kwiatkowski et al[91] used the silyl-linked support (**94**) for oligonucleotide synthesis which was cleaved using 0.5 M TBAF in THF at 20 °C during 34 h or 0.5 M TBAF in DMF at 65 °C during 30 min. This linker was found to be stable towards conc. NH_3 (aq) as well as to acidic solutions including 1% CCl_3CO_2H in CH_2Cl_2 (36 h) and 3% $CHCl_2CO_2H$ in dichloroethane (36 h).



Scheme 1.13 Siloxane-linked supports cleavable under mild conditions.

A novel Silyl-Linked Controlled Pore Glass (SLCPG) support (**95**) was recently reported[92] by the Aston group and further developed as part of this thesis work specifically for the synthesis of oligomers containing base-labile substituents. Crucially, it was found that the support may be cleaved under neutral conditions using a TBAF/AcOH mixture (1:1) in

THF during 60 s (Scheme 1.13). Cleavage of the cyanoethyl phosphorus protecting groups was thereafter achieved by heating the product solution at 55 °C during 2 h. A C^{Bz}- and T-containing sequence was synthesised on the support and was cleaved and deprotected as described.[92] Following purification, the base-labile benzoyl protecting group was still present on the otherwise fully deprotected sequence thus indicating the suitability of this support for the synthesis of sequences continuing base-labile functionality.



96 2

Scheme 1.14 Synthesis of temozlomide (**2**) via the TMS methyl derivative (**96**) using fluoride catalysis.

The speed of support cleavage under mild conditions made SLCPG a suitable alternative to the succinate CPG (**78**) for direct incorporation of the drugs during DNA synthesis. The tetrazinone ring is known to be stable towards prolonged treatment with TBAF/AcOH. During a new synthetic route towards temozolomide (**2**), Wang et al[9] cleaved the trimethylsilyl group of (**96**) using TBAF/AcOH generating **2** (Scheme 1.14).

Chapter 2

Synthesis of Oligonucleotide Conjugates of Temozolomide (Temodal) and Mitozolomide

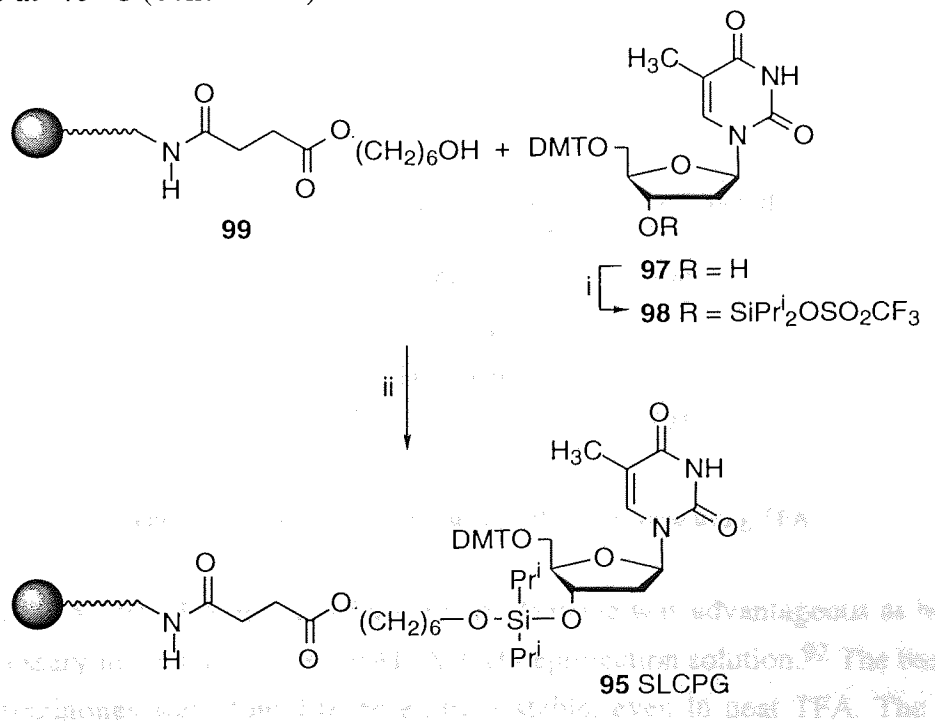
2.1 Introduction

From the literature survey of the various strategies available for potential conjugation of the drugs (**1**) and (**2**) to oligonucleotides (Chapter 1), the following routes were chosen. Phosphoramidite methodology was used to incorporate amino-alkyl tethers at both the 5'-end and at the C4 of a single cytosine base of a sequence to allow post-synthetic support-bound attachment of the drugs. Attempts were made to directly synthesise these same sequences using phosphoramidite methodology. All routes were attempted using the SLCPG support (**95**).

2.2 SLCPG support (95)

2.2.1 Synthesis of DMTdT-derivatised SLCPG support (95)[93]

The DMTdT nucleoside (**97**) was transformed into the known triflate[94, 95] (**98**) by reaction with *bis*-(trifluoromethanesulfonyl)diisopropylsilane in the presence of imidazole at -40 °C (Scheme 2.1).



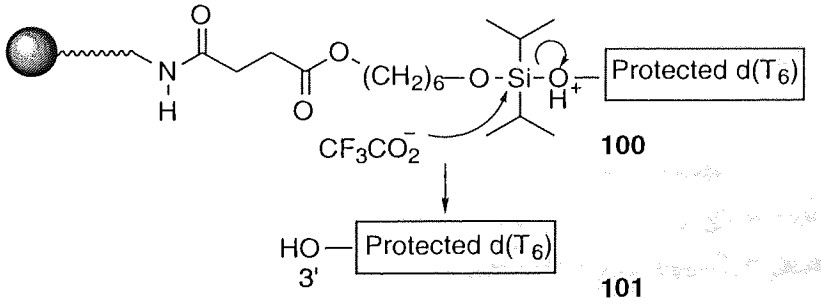
Scheme 2.1 Synthesis of SLCPG support (**95**). i, imidazole, MeCN-DMF (1:1), *bis*-(trifluoromethanesulfonyl)diisopropylsilane, -40 °C, 15 min; ii, 2,6-lutidine, DMAP, r.t., 19 h.

Addition of the reaction mixture containing **98** to the derivatised CPG (**99**)[93] in the presence of 2,6-lutidine and DMAP at r.t., afforded the SLCPG support (**95**) during 19 h. The trityl loading[76] of **95** was determined to be 30 μmol g^{-1}.

2.2.2 New procedures for cleaving SLCPG support (95)

During the course of this work, two new methods for cleaving the SLCPG support (**95**) were developed. The use of $Et_3N.3HF$[96] resulted in substantial support-cleavage (94%) during 15 min via fluoride-catalysed attack at the silicon centre. A number of disadvantages associated with the use of this reagent resulted in its limited use throughout the course of this work. It was found that only 66% of a quantity of CPG could be recovered following treatment with $Et_3N.3HF$ (1 mL) during 10 min, indicating that $Et_3N.3HF$ had dissolved a significant amount of the actual silica support.[80] Removal of $Et_3N.3HF$ by evaporation under low vacuum was problematic due to its high boiling point (78 °C at 1.5 mbar).[96]

A 2% TFA-CH_2Cl_2 solution was found to cleave 98% of the fully protected sequence (**101**) from the support (**95**), during 1 h (Scheme 2.2). The mechanism of support-cleavage was thought to be acid-catalysed, where protonation of an O atom attached to Si afforded intermediate (**100**). Nucleophilic attack by the trifluoroacetate anion at the activated Si centre presumably resulted in detachment of the sequence (**101**) from the support.



Scheme 2.2 Proposed mechanism of SLPCG cleavage using TFA.

The use of the acidic TFA solution for support-cleavage was advantageous as buffering was unnecessary in contrast to the TBAF-AcOH deprotection solution.[92] The base-labile imidazotetrazinones were found to be entirely stable, even in neat TFA. The volatile solution used for cleavage of the support was readily evaporated, under low vacuum, thereby making a time-consuming desalting step unnecessary. Although benzoyl-protected nucleosides, particularly 2'-deoxyadenosine, are known to depurinate under

acidic conditions, McBride et al[97] found that all of the four natural nucleosides are adequately stable in 2% TFA-H_2O solution for periods of up to 1 h. The 2% TFA-CH_2Cl_2 solution was susequently used for conjugate synthesis.

2.3 Post-synthetic, support-bound conjugation of drugs (1) and (2) to oligonucleotides at the 5'-end.

2.3.1 Considerations for support-bound conjugate synthesis

5'-d(TTTCTTTTCTCTCTT) **27**

In order to be able to couple the drugs to the support-bound oligonucleotide (**27**) (Section 1.5.4), it was noted that the base-labile protecting groups could not be cleaved from the support-bound sequence once either of the base-sensitive drugs (**1**) or (**2**) had been introduced into the sequence. Therefore, it was necessary to couple the drugs to a fully deprotected sequence bound to SLCPG support. The following points were addressed:

1. The standard benzoyl base-protecting group for C was not suitable as the basic conditions necessary for its removal would have caused cleavage of the base-labile succinate linkage of the SLCPG support and it was this point that was addressed initially.

2. An SLCPG-compatible protecting group was required for the 5'-end amino group.

2.3.2 *t*-Butylphenoxyacetyl (BPA^t) DNA base protecting group

Polushin et al[81] investigated the use of BPA^t as a base protecting group for the exocyclic amino functions of A, C and G. Complete removal of the cyanoethyl protecting groups from phosphorus and the BPA^t groups was achieved by treating sequences with ethanolamine during 10 min. Crucially, Polushin et al found that the succinate linkage was cleaved more slowly, during 3 h in ethanolamine. Therefore, provided that the succinate linkage of SLCPG (**95**) was not extensively cleaved by ethanolamine during 10 min, the use of the commercially available, BPA^t-protected C phosphoramidite provided a means of synthesising the fully deprotected sequence (**27**), bound to SLCPG support.

2.3.3 Stability of SLCPG support (95) towards ethanolamine

To assess the stability of SLCPG support (**95**) towards ethanolamine, quantities of the DMTdT-derivatised support (**95**) were treated with ethanolamine for periods up to 24 h. The trityl load of these supports was then determined and compared with the load of the untreated support. After 20 min only 6% of the support was cleaved, but after 100 min, a substantial loss of 15% of the support was observed (Figure 2.1). Reduction in loadings of the treated supports was attributed to base-catalysed cleavage of the succinate linkage of the SLCPG support by the treatment with ethanolamine.

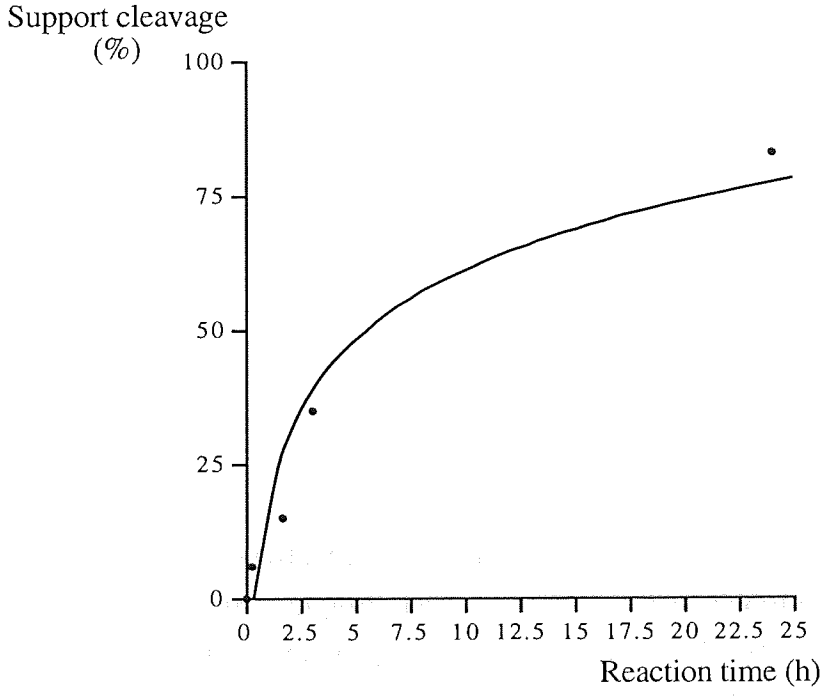


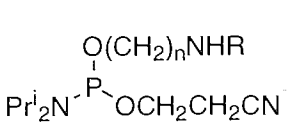
Figure 2.1 Time-dependant cleavage of the succinate linkage of SLCPG support (**95**) using ethanolamine.

Thus, ethanolamine deprotection was considered a viable route towards synthesis of deprotected support-bound sequences. It remained to find a suitably-protected amino-linker phosphoramidite to allow 5'-end modification.

2.3.4 Choice of phosphoramidite for 5'-end derivatisation

Functionalisation of the 5'-end of a sequence with phosphoramidite (**48**) was undesirable as removal of the base-labile, trifluoroacetyl protecting group would have caused cleavage of the succinate linkage of the SLCPG support. However, the acid labile MMT protecting group of **47**, was considered more suitable for this work. A variety of acidic conditions have been used to cleave MMT-amino protecting groups. These have

included treatment with 80% AcOH-H_2O (1 h[98]), 3% $CHCl_2CO_2H$-CH_2Cl_2 (5 min[99], or 15 min[100]), 3% $CHCl_2CO_2H$-CH_2ClCH_2Cl (10 min[101]) and 3% CCl_3CO_2H-CH_2Cl_2 (3 min[102]).



47 R = MMT, n = 6
48 R = $CF_3C(O)$, n = 6
102 R = DMT, n = 6
103 R = DMT, n = 5

The 3% $CHCl_2CO_2H$-CH_2Cl_2 solution was the most suitable reagent for cleaving the MMT group since it provided rapid removal of this group and it was readily available as the Deblock solution on the DNA synthesiser. Also, it is known that the disiloxane linkage of the SLCPG support is stable towards 3% $CHCl_2CO_2H$-CH_2Cl_2 during the acidic Deblock phase of DNA synthesis[92]. However, it was necessary to ensure that the support was stable under those conditions for the time period necessary to cause cleavage of the MMT protecting group.

2.3.5 Stability of SLCPG towards 3% $CHCl_2CO_2H$-CH_2Cl_2 solution

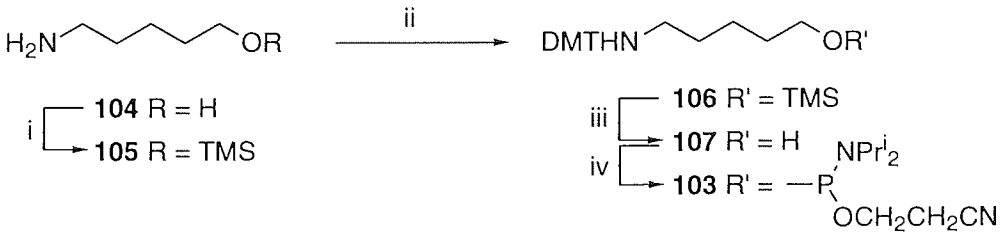
A quantity of SLCPG support (**95**) (28 μmol g^{-1}) was treated with 3% $CHCl_2CO_2H$-CH_2Cl_2 solution during 10 min. As the DMT group was completely cleaved under these acidic conditions, it was necessary to determine the load of dT nucleoside remaining on the support. The support was washed and cleaved using TBAF-AcOH during 1 min to remove dT nucleoside from the support. The support load was determined to be 15.6 μmol g^{-1}, a loss of 44% of the support, presumably through acid-catalysed cleavage of the disiloxane linkage. The necessary conditions required for cleavage of the MMT group were therefore incompatible with the SLCPG support. A more acid-labile amino-protecting group was sought.

2.3.6 The DMT amino-protecting group

Phosphoramidite (**102**) containing a DMT amino-protecting group was previously prepared by Sinha et al[103] The enhanced acid-lability of the DMT group compared with the MMT group made it a suitable candidate for this work, with the added advantage that it could be removed automatically at the end of DNA synthesis. Molecular modelling studies revealed that, at the 5'-end, optimal positioning of the drug would be achieved using phosphoramidite (**103**) where n = 5 (Scheme 2.3).

Compound (**103**) was prepared according to Sinha et al (Scheme 2.3). 5-Aminopentan-1-ol (**104**) was reacted with trimethylsilyl chloride (TMSCl) at 0 °C affording **105**.

Material (**105**) was then reacted with DMTCl affording **107** in 57% yield following cleavage of the TMS group of **106** using MeOH. Phosphitylation of the hydroxyl of **107** using 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite during 45 min afforded the product (**103**) in 88% yield following purification by flash chromatography.

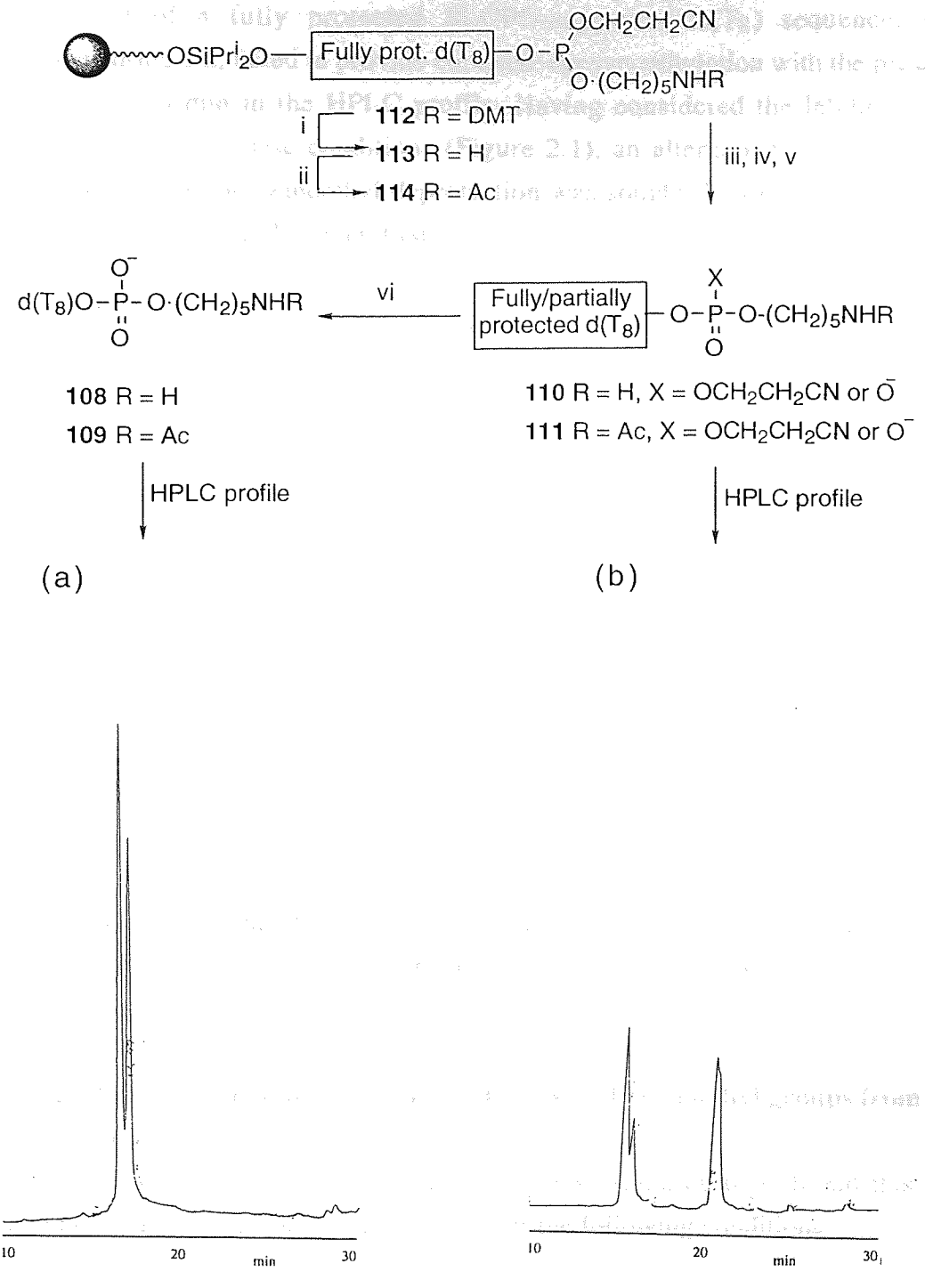


Scheme 2.3 Synthesis of phosphoramidite (**103**). i, TMSCl, pyridine, 0 °C; ii, DMTCl, r.t., 12 h; iii, MeOH; iv, DIPEA, 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite, THF, 45 min, r.t.

2.3.7 Solid-supported DNA synthesis of sequence (108) using phosphoramidite (103)

Sequence (**108**) was chosen as a simple model target (Scheme 2.4). It was synthesised on normal CPG using phosphoramidite (**103**). HPLC analysis of the product mixture revealed two peaks with t_Rs 14.8 and 15.5 min were identified by ESMS as sequences (**108**) with MW 2535.3 (calcd 2535.5) and (**109**) with MW 2578.2 (calcd 2578.5), respectively. The sequence (**108**) was synthesised in 76% overall yield, based on trityl yield and a coupling efficiency of 83% was observed for the addition of phosphoramidite (**103**). Sequence (**108**) was synthesised on SLCPG support and a portion of the support-bound sequence was treated with ethanolamine during 10 min to remove the cyanoethyl protecting groups. Standard fluoride-catalysed cleavage of the support afforded the crude oligonucleotide mixture. The HPLC profile displayed a number of large peaks with t_Rs of 15.9, 16.3, 20.5 and 21.0 min (Scheme 2.4 (b)).

A portion of this oligonucleotide mixture was treated with conc. NH_3 (aq) at 55 °C during 12 h. HPLC analysis of the resulting mixture (Scheme 2.4 (a)) indicated the presence of only two peaks with t_Rs 17.0, 17.5 min which were identified as sequences (**108**) and (**109**) respectively, by co-injection with authentic samples. Disappearance of the peaks with t_R ~ 20 min following treatment with conc. NH_3 (aq) indicated that these peaks were due to partially and/or fully cyanoethyl-protected sequences (**110**) and (**111**). Repeated attempts to deprotect sequence (**108**) while support-bound using ethanolamine during 10 min failed. Longer exposures to ethanolamine (15, 20 and 40 min) failed to provide the fully deprotected sequence.



Scheme 2.4 Synthesis of sequence (**108**) and by-product (**109**). HPLC profiles of the crude product mixture following deprotection with (a) ethanolamine and subsequently (b) conc. NH_3 (aq). i, 1*H*-tetrazole (DNA Coupling step); ii, Ac_2O, imidazole (DNA CAP step), THF; iii, completion of DNA synthesis; iv, ethanolamine, r.t., 10 min; v, TBAF (0.5 M)-AcOH (0.5 M), THF, 60 s; vi, conc. NH_3 (aq) at 55 °C, 12 h.

Even treatment of a fully protected SLCPG-supported d(T_8) sequence, with ethanolamine during 3 h, failed to provide complete decyanoethylation with the presence of peaks at t_R ~ 20 min in the HPLC profile. Having considered the lability of the succinate linkage under these conditions (Figure 2.1), an alternative set of SLCPG-compatible conditions for cyanoethyl deprotection was sought. Firstly, an explanation for the formation of **109** will be given first.

2.3.8 Mechanism of formation of by-product sequence (109)

Following coupling of phosphoramidite (**103**) to the support-bound oligomer forming **112** during the final coupling step of DNA synthesis,[104] the acidic activating agent, tetrazole (pK_a = 4.8), caused detritylation of some of the DMT-amino groups of **112** forming **113** (Scheme 2.4).[105] The free amino group of **113** was then acetylated during the proceeding CAP phase of the DNA synthesis cycle by reaction with acetic anhydride forming **114**, which upon completion of DNA synthesis, and complete deprotection and cleavage afforded **109**. DNA synthesis using standard phosphoramidites is affected to a lesser extent by this process since the 5'-DMT-O protecting group is more acid-stable than the DMT-NH protecting group of **112** since the oxygen atom is substantially less basic than the nitrogen atom. Although 46% of the product sequence (**108**) was lost due to formation of sequence (**109**), based on peak area of Scheme 2.4 (a), the quantity of **108** formed was sufficient to allow synthesis, purification and evaluation of the desired conjuagtes. 4,5-Dicyanoimidazole[106] (DCI) has recently been made available commercially as a replacement for tetrazole during DNA synthesis. The weak acidity (pK_a = 5.2)[105] of DCI relative to tetrazole would benefit future use of phosphoramidite (**103**) for the synthesis of modified DNA sequences.

2.3.9 Reagents other than ethanolamine for removal of cyanoethyl groups from phosphorus

During early work with cyanoethyl phosphoramidites, Sinha et al[107] found that the cyanoethyl protecting group could be cleaved under the following conditions:

1. 10% Bu^tNH_2-pyridine, 15 min.

2. 20% Et_3N-pyridine, 2 h.

Alul et al[79] also found that anhydrous diisopropylamine (DIPA) was effective in decyanoethylation of the phosphotriesters of an eleven base, support-bound DNA sequence during 14 h without cleaving the highly base-labile oxalyl-linked support (**79**)

(Figure1.8). The effectiveness of all of these reagents at removing the cyanoethyl protecting groups from sequence (**108**) while fully protected and support-bound was assessed. Portions of the fully cyanoethyl-protected support-bound sequence (**108**) were treated with these reagents using a variety of temperatures and reaction times. Sequences were then cleaved from the support using TBAF as fluoride source, affording crude oligonucleotides which were analysed by HPLC. The individual conditions found to give complete decyanoethylation were as follows:

1. Anhydrous DIPA at 60 °C during 17 h.

2. Pyridine-Et_3N-(4:1) during 24 h.

3. Anhydrous Bu^tNH_2 during 15.5 h. (In this case, 2% TFA-CH_2Cl_2 was used to cleave the deprotected sequence from the support)

Of these, Bu^tNH_2 was selected as the optimal decyanoethylating agent since it provided fast, efficient decyanoethylation at ambient temperature as well as being cheap and available commercially in anhydrous form. Since the reagent was used directly, neat, the preparation of a stock solution was unnecessary.

To test that the SLCPG support was stable towards Bu^tNH_2, two portions of SLCPG (**95**) (53.3 μmol g^{-1} load of DMTdT) were treated with anhydrous Bu^tNH_2 during 15.5 h. The support was washed and the average trityl load was determined to be 53.7 μmol g^{-1}. Comparison of the support loadings before and after treatment with Bu^tNH_2 indicated that the SLCPG support was stable towards this reagent during 15.5 h. It remained to determine whether a mixed C^{BPA^t} and T containing sequence could be completely deprotected, using a combination of Bu^tNH_2 and ethanolamine, while still bound to SLCPG support.

2.3.10 Synthesis of sequence 5'-d(TCCTCTCT) (115)

5'-d(TCCTCTCT) **115**

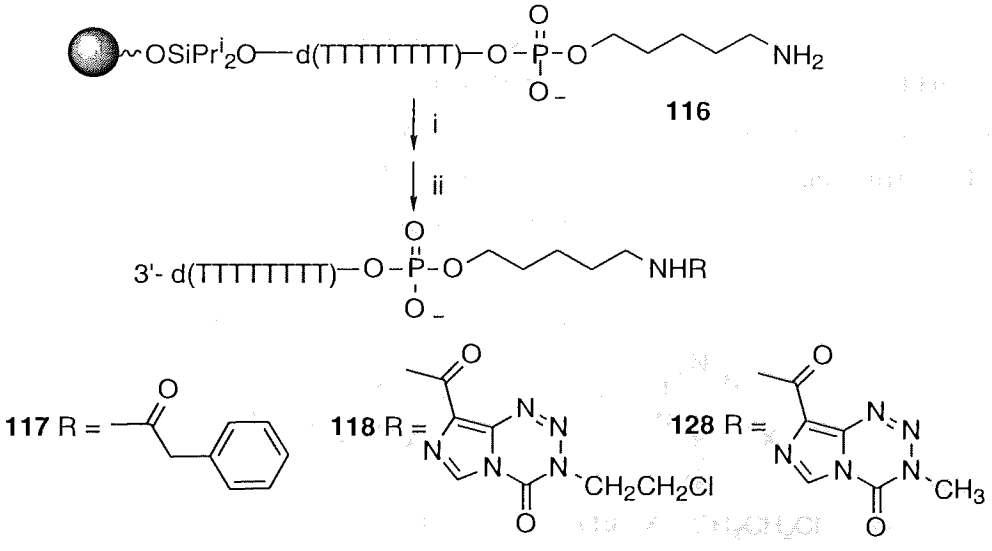
Sequence (**115**) was chosen as a model sequence with which to validate the Bu^tNH_2 - ethanolamine deprotection procedure. **115** was synthesised, initially on normal CPG, using the C^{BPA^t} phosphoramidite monomer. Following cleavage and deprotection using conc. NH_3 (aq) during 16 h, the crude sequence was purified by HPLC (t_R 13.7 min) and

its identity confirmed by ESMS analysis (MW calcd 2310.4; found 2311.9). Sequence (**115**) was also synthesised on SLCPG support which was then treated with Bu^tNH_2 overnight and subsequently with ethanolamine during 30 min. The fully deprotected sequence was cleaved from the washed support using $Et_3N.3HF$ during 15 min. The HPLC profile displayed one main peak, the identity of which was confirmed as **115** by co-injection with an authentic sample. The deprotection procedure was therefore successful. Solid-phase preparation of a C- and T-containing drug-conjugate sequence was thus synthetically feasible.

2.3.11 Conjugation of drugs (1) and (2) to sequence (108)

In order to assess and optimise drug-coupling conditions, it was decided to initially use the sequence (**108**) for drug conjugate formation in order to minimise the potential for side reactions with the exocyclic amino funtions of C. PyBOP[108] was selected as coupling agent as it has been used successfully by other workers[109] to carry out similar types of amide bond-forming reaction. Phenylacetic acid was chosen as a robust model compound to be used initially in place of the drugs for coupling to the SLCPG support-bound sequence (**116**).

2.3.12 Synthesis of the phenylacetyl conjugate sequence (117)



Scheme 2.5 Synthesis of conjugate sequences (**117**), (**118**) and (**128**). i, ROH, DIPEA, PyBOP, DMF, r.t., 2 h; ii, 2% TFA-CH_2Cl_2, r.t., 1 h.

The fully deprotected, support-bound sequence (**116**) (Scheme 2.5) (1 eq) was synthesised and deprotected using Bu^tNH_2 as described (Section 2.3.9). A solution of

DMF containing phenylacetic acid (1200 eq), PyBOP and DIPEA was prepared and added to the SLCPG-bound sequence (**116**) (Scheme 2.5). After 2 h reaction, the support was washed with DMF, MeCN and ether and the oligonucleotide was cleaved from the support using 2% TFA-CH_2Cl_2 (1 h). The product was purified by HPLC (t_R 16.6 min) (Figure 2.2) and its identity confirmed as **117** by ESMS analysis (MW calcd 2653.5, found 2654.0), thus validating the derivatisation procedure.

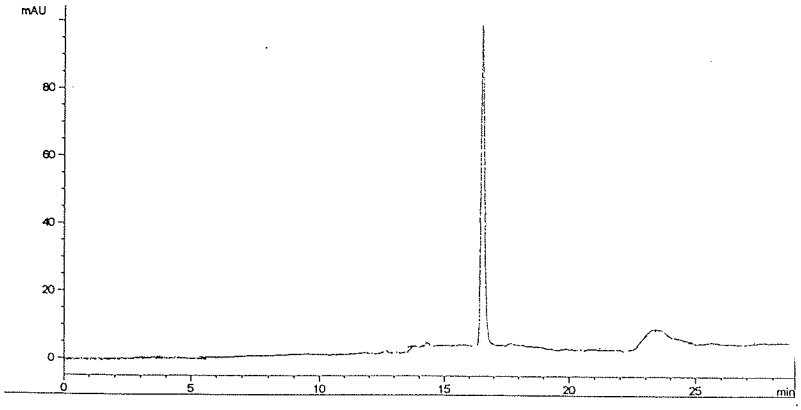
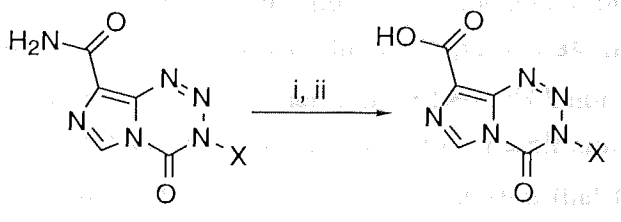


Figure 2.2 HPLC profile of the phenylacetyl conjugate sequence (**117**).

2.3.13 Synthesis of the mitozolomide conjugate (118)

For conjugation to the amino function of the support-bound sequence (**116**), suitable derivatives of both **1** and **2** were synthesised according to a literature procedure.[110] Nitrous acid hydrolysis of **1** and **2**, afforded the 8-carboxylate derivatives (**119**) and (**120**) in yields of 95% and 80%, respectively (Scheme 2.6).



1 X = CH_2CH_2Cl **119** X = CH_2CH_2Cl

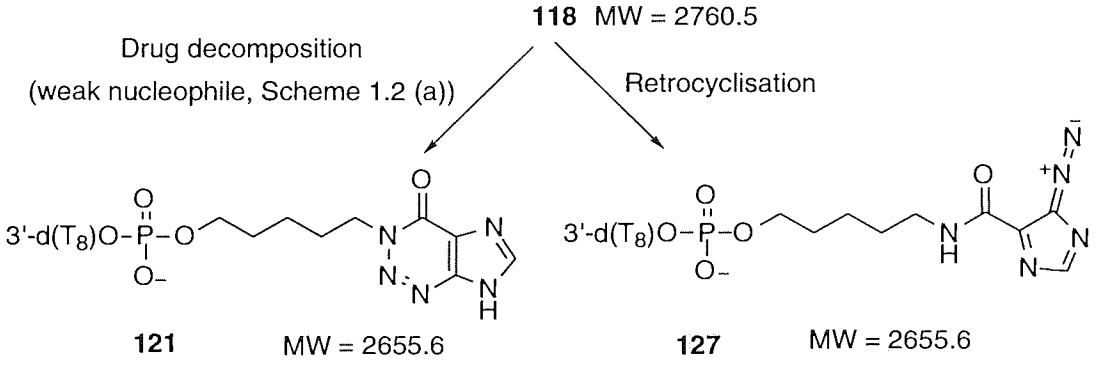
2 X = CH_3 **120** X = CH_3

Scheme 2.6 Synthesis of **119** and **120**. i, H_2SO_4; ii, $NaNO_2$, H_2O.

A quantity of the deprotected support-bound sequence (**116**) (1 eq) was treated with a solution of DMF containing **119** (385 eq), PyBOP and DIPEA during 2 h. The product

(Scheme 2.5) was cleaved from the washed support using 2% TFA-CH_2Cl_2 solution (1 h) and the solvent evaporated. The HPLC profile of the crude sequence is shown in Figure 2.3 (a).

Peak A (t_R 13.6 min) (Figure 2.3 (a)) was isolated by semi-preparative HPLC and its structure was confirmed as **109** by ESMS analysis (MW calcd 2578.5, found 2580.0). Peak B (t_R 14.1 min) was thought to be an impurity from the coupling reaction. Peak C (t_R 16.6 min) was purified (Figure 2.3 (b)) by HPLC and analysed by ESMS which indicated that the sequence had the incorrect molecular weight (2656.9). The calculated molecular weight for the drug conjugate (**118**) was 2760.5. It was thought that the drug function of conjugate (**118**) had decomposed[7] to perhaps produce the azahypoxanthine conjugate (**121**) (Scheme 2.7) (MW calcd 2655.6) via the route outlined in Scheme 1.2 (a): ring-opening of the tetrazinone ring followed by intramolecular cyclisation.



Scheme 2.7 Decomposition routes of the mitozolomide conjugate (**118**).

However, UV analysis of the conjugate after ESMS analysis showed a λ_{max} at 265 nm, consistent with the absorption due to the oligonucleotide and a much weaker, maximum at λ_{max} 329.6 (Figure 2.3 (f)). Curiously, this absorption was indicative[7] of the intact tetrazinone ring-system of the drug. Neither unmodified DNA nor 2-azahypoxanthine (**8**) (λ_{max} 250, 288 nm at pH 7.4)[111] exhibit an absortion maximum in the 300 to 350 nm region of the UV spectrum. Thus, UV analysis suggested that the isolated conjugate contained the intact drug. Further evidence to support this was obtained following an experiment in which the isolated conjugate was treated with conc. NH_3 (aq) at 45 °C during 1 h, conditions under which the azahypoxanthine conjugate (**121**) should be stable. UV analysis of the resulting sample did not show the previous absorption in the 300-350 nm region (Figure 2.3 (g)). The HPLC profile of the sample shown in Figure 2.3 (c) indicated the presence of one major peak (t_R 14.1 min) and a number of low abundance peaks, all with shorter retention times (t_R 11.5 to 14.0 min) compared with

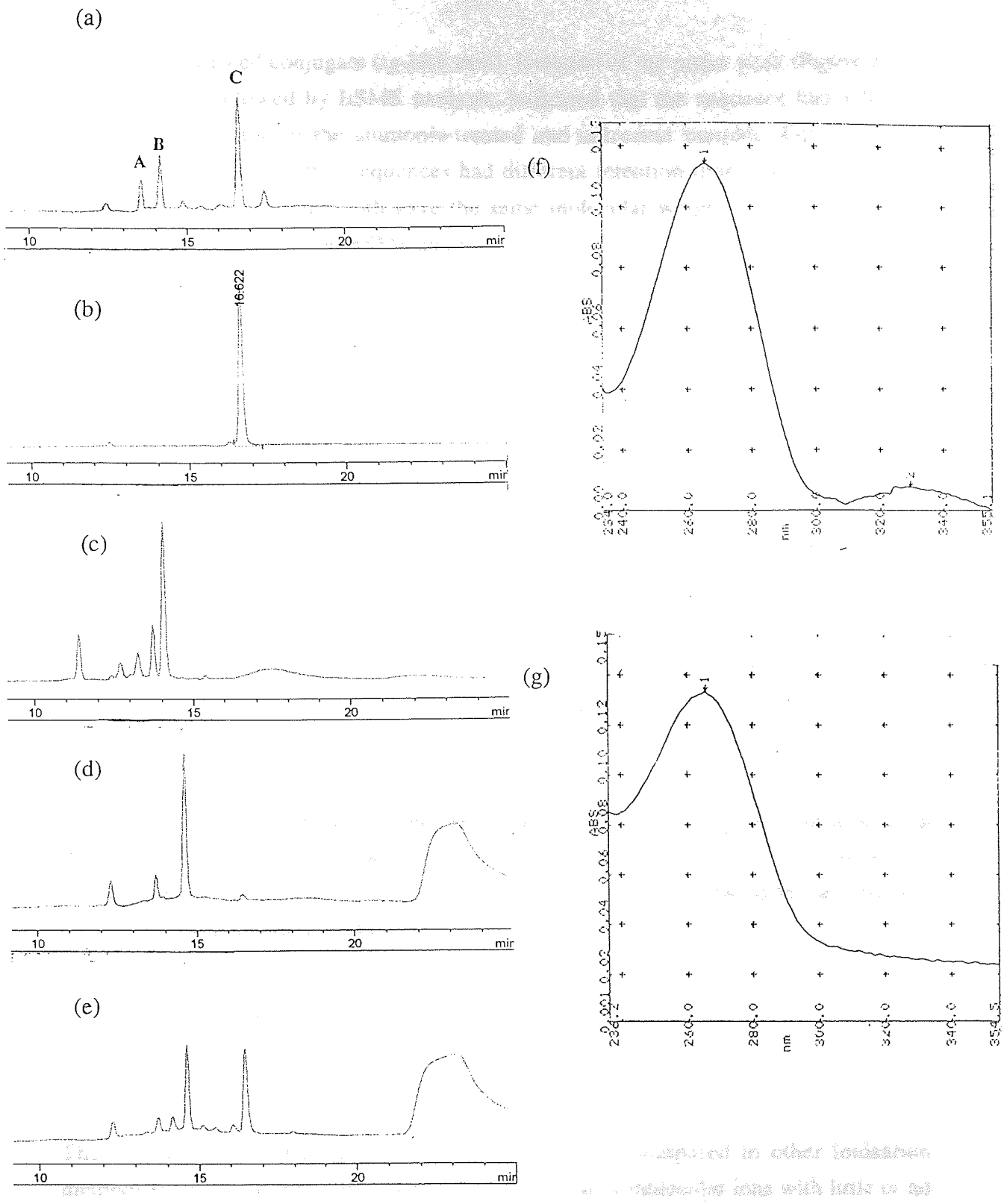


Figure 2.3 HPLC profile of (a) crude oligonucleotide (**118**); (b) oligonucleotide (**118**) after purification; (c) crude mixture resulting from treatment of (b) with conc. NH_3; (d) purification of major peak in (c), (e) coinjection of (b) and (d). UV analysis of oligonucleotide (**118**) (f) prior to treatment with conc NH_3 and (g) following treatment with conc. NH_3 (aq).

that of the untreated conjugate (t_R 16.6 min). Isolation of the major peak (Figure 2.3 (d) t_R 14.6 min), followed by ESMS analysis, indicated that the sequence had a MW of 2656.5. Co-injection of the ammonia-treated and untreated samples (Figure 2.3 (e)) showed clearly that the two sequences had different retention times (t_R 14.4, 16.5 min respectively), even though both gave the same molecular weight (*m/z* 2656.5, 2656.9 respectively) in negative ion ESMS analysis.

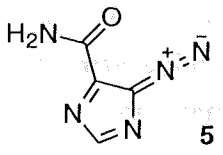
The following evidence suggested that the isolated conjugate was **118**:

1. Broad absorption from 300 to 350 nm region of UV spectrum (Figure 2.3 (f)) consistent with intact drug.

2. The isolated conjugate decomposes in conc. NH_3 (aq) with loss of 300 to 350 nm absorption band whereas 2-azahypoxanthine (**8**) is known to be stable in NH_3 (aq) (Section 1.2.2).

2.3.14 Collision-induced dissociation (CID)

It is generally well known[7] that successful mass spectrometric analysis of the parent unconjugated imidazotetrazinones is troublesome. On the few occasions when molecular ions have been observed for these compounds, the abundance is extremely low and daughter ions are intense, with *m/z* 137 usually the base peak. Stevens et al[7] synthesised a number of imidazotetrazinone analogues with a variety of alkyl and aryl substituents at the N3 position. A major ion (*m/z* 137) was always present in the electron impact (EI)-promoted mass spectra, for all analogues. The same ion was attributed to the thermally stable retro-cycloaddition product (**5**).

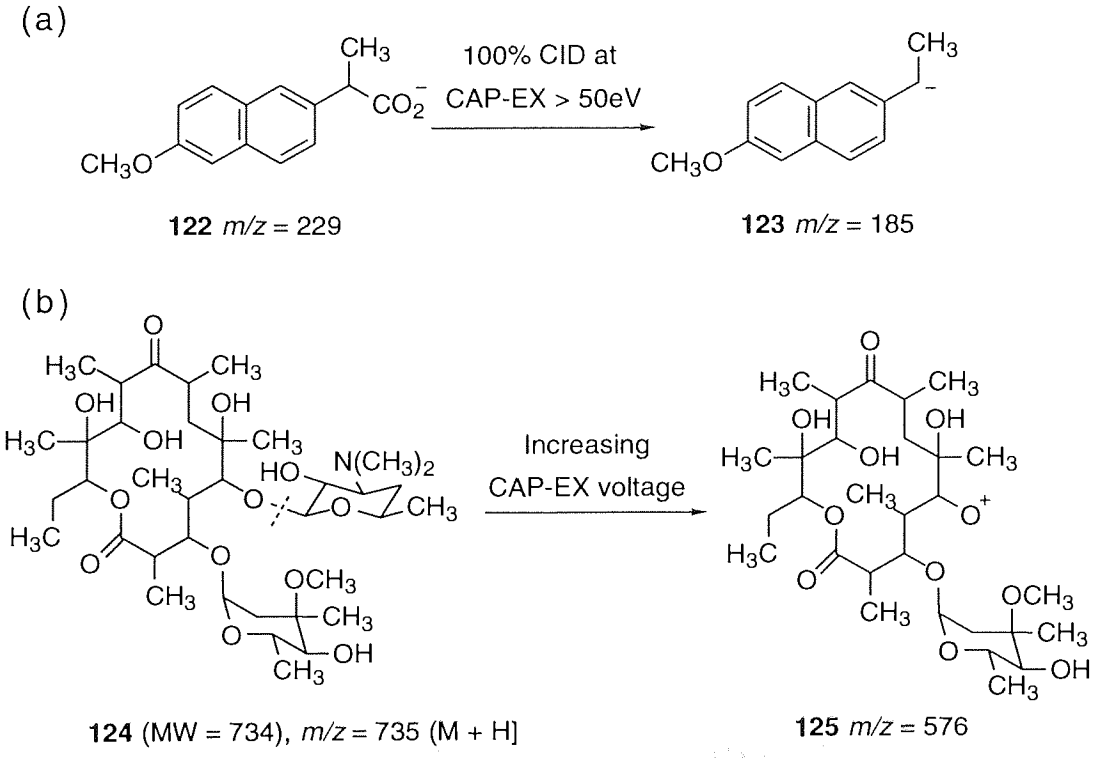


5

The relatively 'soft' nature[112, 113] of the ES process compared to other ionisation methods such as electron impact (EI) allows detection of molecular ions with little or no fragmentation. However, a process known as collision-induced dissociation (CID) can occur during ESMS when a voltage, known as the CAP-EX voltage, across a region through which analyte ions pass on route to the detector, is sufficiently high. CID explains the fragmentation of analytes when ions are accelerated into collisions with molecules of drying gas, as a result of the voltage differential. CID is a technique

commonly used to provide structural information about particular analytes.[112, 114] The amount of fragmentation can be quite extensive for certain types of molecules particularly those which fragment to produce particularly stable species.

2.3.15 Literature precendents of CID



Scheme 2.8 (a) CID-induced decarboxylation of naproxen (**122**) and (b) CID-induced fragmentation of erythromycin (**124**).

Naproxen[114] (**122**) is a drug which undergoes fragmentation due to CID during ESMS (Scheme 2.8 (a)). **122** undergoes decarboxylation forming **123** (*m/z* 185) under conditions (CAP-EX > 50 eV) where it is considered most samples ions would survive. At low collision energy (CAP-EX < 50 eV), the molecular ion *m/z* 229 of the intact species (**122**) is observed as the only peak in the spectrum.

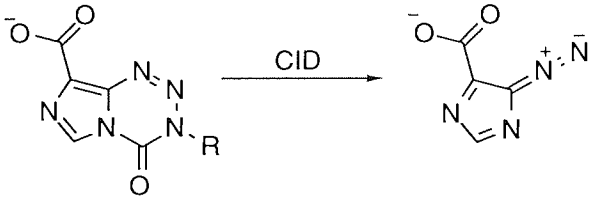
At low CAP-EX voltage, ESMS analysis of erythromycin[112] (**124**) shows a predominant ion at *m/z* 735 corresponding to the protonated molecular ion (M + H) (Scheme 2.8 (b)). When analysed at a higher CAP-EX voltage, the *m/z* 735 is no longer the base peak, while two dominant ions are observed at *m/z* 576 and 158. These ions are consistent with

loss of the amino sugar group from erythromycin to give **125**, further demonstrating the effect of the CAP-EX voltage on molecular ion abundance.

The absence of a molecular ion consistent with conjugate (**118**) was therefore thought to be a result of the ES process itself. To test whether conjugate (**118**) was undergoing CID during ESMS analysis, the acid derivatives (**119**) and (**120**) of the parent drugs were analysed by ESMS.

2.3.16 ESMS analysis of the drug-derivatives (119) and (120)

Equimolar solutions (28.2 μM) of **119** and **120** were prepared and analysed directly by ESMS under identical conditions to those used for analysis of the oligonucleotide conjugates, that is, using a CAP-EX volatge which cycled from 150-300 eV during analysis.



119 R = CH_2CH_2Cl *m/z* = 242 $(M-1)^-$ **126** *m/z* = 137

120 R = CH_3 *m/z* = 194 $(M-1)^-$

Scheme 2.9 CID-induced retrocyclisation of **119** and **120**.

Table 2.1 Effect of CAP-EX voltage on the stability of imidazotetrazinone ring-system.

CAP-EX (eV)	Ions present (*m/z*), abundance (%)	
	119	**120**
150 to 300	242 (0%), 137 (100%)	194 (1.7%), 137 (100%)
10	242 (0%), 137 (100%)	194 (24%), 137 (100%)

A molecular ion at *m/z* 242 was not observed for **119** but a single major ion was observed at *m/z* 137 consistent with the retro-cycloaddition fragment 5-diazoimidazole-4-carboxylate (**126**) (Table 2.1, Scheme 2.9). For **120**, a weakly abundant molecular ion was observed (M - H, *m/z* 194, 1.7%) consistent with intact **120**, but more importantly, the base peak was at *m/z* 137 which is consistent with **126** (Table 2.1). It was found that when the samples were run at a lower, constant CAP-EX voltage (10 eV), an increase in

the abundance of the molecular ion *m/z* 194 (24%) of **120** was observed thus confirming that CID was causing retrocyclisation of the drugs. No molecular ion was observed for **119** under these conditions.

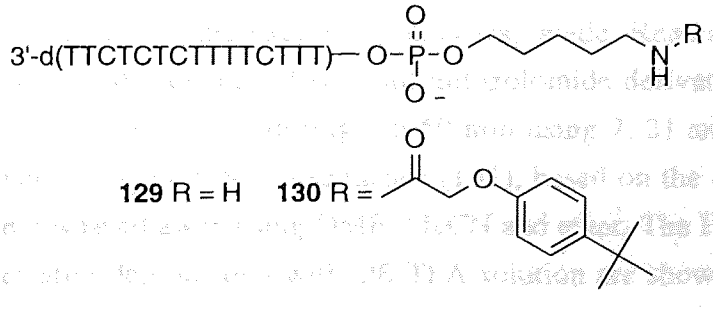
These results confirmed that the isolated conjugate was the desired derivative (**118**) with ESMS indictating the presence of the retrocyclisation conjugate (**127**) (Scheme 2.7).[115] Presumably, treatment of **118** with conc. NH_3 (aq) resulted in formation of the azahypoxanthine conjugate (**121**) (Scheme 2.7), as indicated by ESMS, via the mechanism detailed in Scheme 1.1.

2.3.17 Synthesis of the temozolomide conjugate (128)

The temozolomide conjugate (**128**) (Scheme 2.5) was prepared[115] by a similar method to that used to prepare **118** using the temozolomide derivative (**120**). The crude sequence was HPLC purified (t_R 14.4 min) and its identity was confirmed by ESMS which indicated the presence of the retrocyclisation conjugate (**127**) (Scheme 2.7) (MW calcd 2655.6, found 2656.0). UV analysis of the purified product showed a spectrum very similar to that seen for conjugate (**118**) (Figure 2.3 (f)), where the presence of λ_{max} 329.0 nm suggested that the drug was intact. Having developed an appropriate route for oligonucleotide-drug conjugate synthesis, synthesis of conjugates of sequence (**27**), suitable for targeting duplex DNA, was commenced.

2.3.18 Synthesis of the 5'-end functionalised TFO (129)

Sequence (**129**) was synthesised on SLCPG support in DMT-OFF mode using phosphoramidite (**103**). The sequence was deprotected by treatement with Bu^tNH_2 overnight and with ethanolamine during 30 min. After cleavage from the support using 2% TFA-CH_2Cl_2 (1 h), HPLC analysis showed the presence of a single major peak with an unexpectedly long retention time (25.8 min). ESMS analysis gave *m/z* 4796.0 consistent with that calculated for **130** (4793.9).



129 R = H **130** R =

Since there was no evidence of the expected amino-linker sequence (**129**), it was presumed that a combination of the following led to the formation of **130**.

(a) Amino functions exposed during DNA synthesis as a result of tetrazole-catalysed cleavage of the DMT-N groups were BPA^t-protected by reaction with the BPA^t anhydride CAP solution or by a transamidation reaction involving the base-labile BPA^t-protected C residues forming **130** (Section 2.3.8).

(b) Amino functions exposed at the end of DNA synthesis were BPA^t-protected as a result of the transamidation reaction described above.

Others have observed the transamidation of base-labile protecting groups from nucleotide residues to free primary amine functions within oligonucleotides.[116,117] The sequence (**129**) was synthesised in the DMT-ON mode with omission of the final capping step. The DMT group was only cleaved after all other protecting groups were removed (cyanoethyl, BPA^t). The fully deprotected SLCPG-bound sequence was washed with Bu^tNH_2 and MeCN to remove any residual traces of acid. The desired sequence (**129**) was isolated following HPLC purification in 6.4% yield based upon support loading (MW calcd 4603.8, found 4606.5). The phosphoramidite (**103**) was observed to couple in 60 % yield during DNA synthesis. Sequence (**130**) was still present (7.5% yield) in the HPLC profile and its presence was explained by the transamidation reaction described in (a) above. Even with a loss of 54% of product (**129**) due to the formation of **130** (based on HPLC analysis), sufficient quantities of the drug-conjugates could be isolated for triplexation studies.

2.3.19 Synthesis of mitozolomide and temozolomide conjugates (132) and (134): effect of excess reagents

The problem of unwanted side-reactions, such as those involving exocyclic amino functions of bases, during DNA-derivative synthesis is well known[118] and so an investigation of the effect of using excess reagents was made. Reactions of the fully deprotected, support-bound sequence (**131**) with mitozolomide derivative (**119**) in the presence of PyBOP were carried out during 1 h 50 min using 7, 21 and 76 equivalent excesses of **119** over the support-bound sequence (**131**), based on the support loading. Excess reagents were washed away using DMF, MeCN and ether. The HPLC profiles of the crude sequences after deprotection with 2% TFA solution are shown in Figures 2.4 (a), (b) and (c).

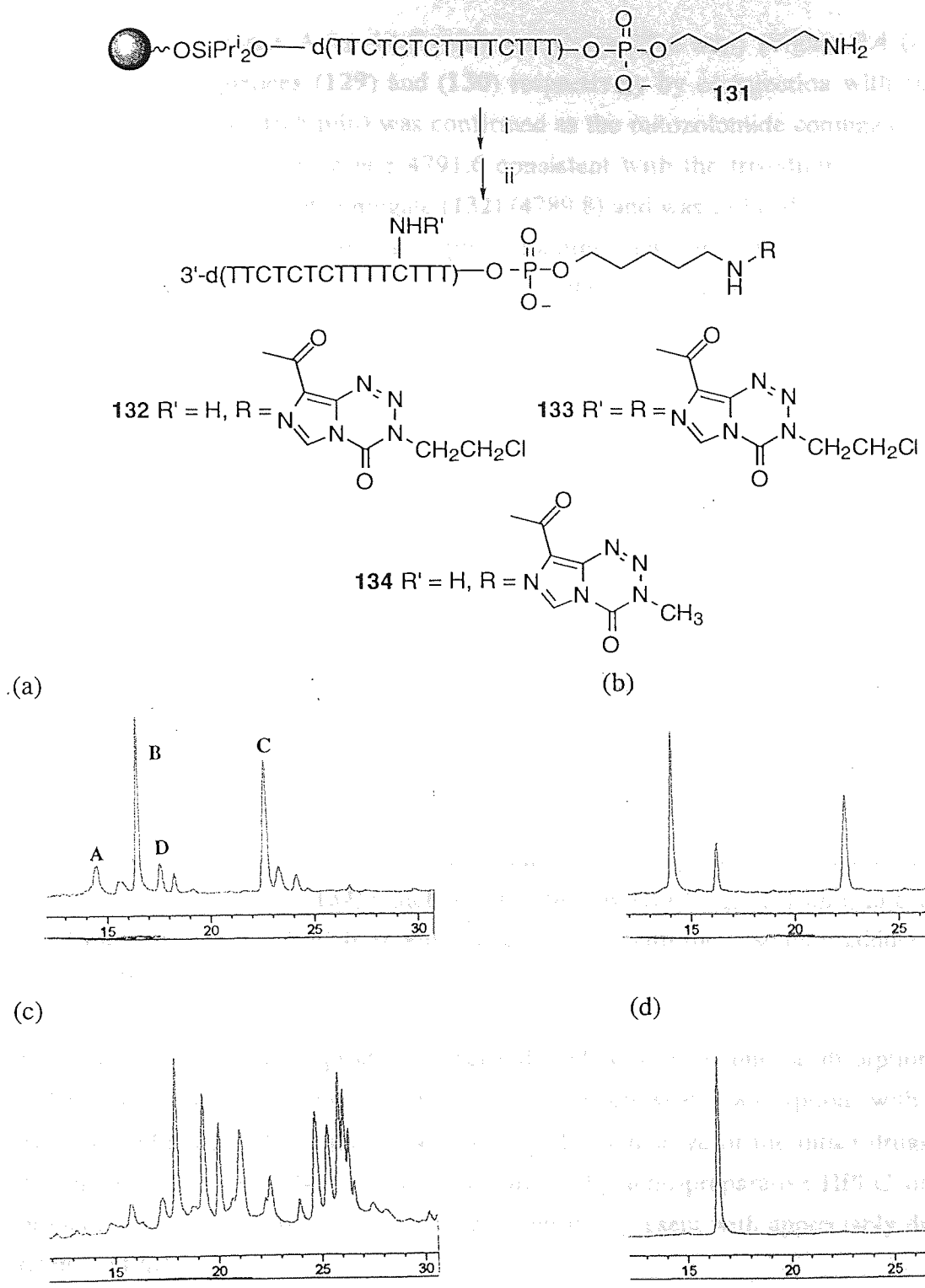


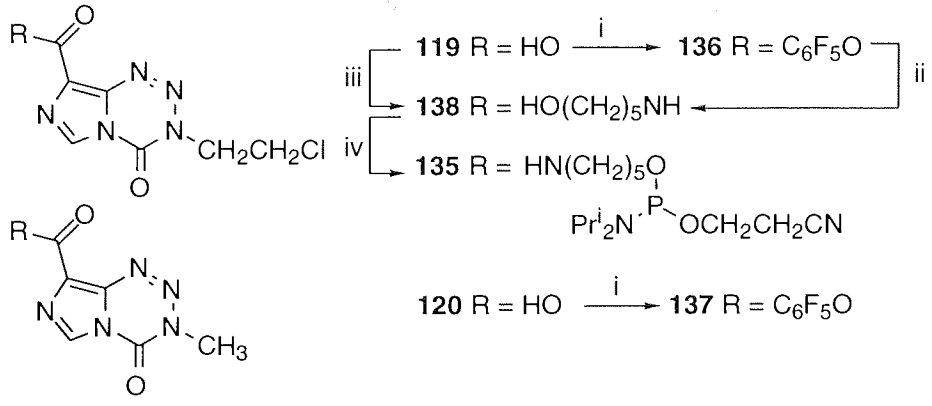
Figure 2.4 Synthesis of conjugates (**132**), (**133**) and (**134**). i, ROH, DIPEA, PyBOP, DMF, r.t., 1 h 50 min; ii, 2% TFA-CH_2Cl_2, r.t., 1 h. HPLC profiles of the crude oligonucleotide mixtures obtained following reaction of **131** with (a) 21 eq, (b) 7 eq and (c) 76 eq of **119**; (d) HPLC profile of the pure conjugate (**132**).

The identity of peaks A (t_R 14.5 min) and C (t_R 22.6 min) (Figure 2.4 (a)) were confirmed as sequences (**129**) and (**130**) respectively by co-injection with authentic samples. Peak B (t_R 16.5 min) was confirmed as the mitozolomide conjugate (**132**) by ESMS analysis which gave *m/z* 4791.6 consistent with the trisodium adduct of the retrocycloaddition form of conjugate (**132**) (4789.8) and was isolated (Figure 2.4 (d)) in 4.5% yield based on the initial support loading. Two groups of peaks of lower abundance than the product were observed to elute with slightly longer t_R than peaks B and C, respectively. In the case where 7 eq of **119** were used for coupling (Figure 2.4 (b)), a small amount of product (**132**) (t_R 16.2 min) was observed as well as substantial amounts of **129** and **130** (t_R 14.0, 22.5 min respectively). No other significant peaks were observed in the HPLC profile. As the number of equivalents of **119** used in the reaction was increased to 76, it was clear from Figure 2.4 (c) that other major peaks were present compared with Figure 2.4 (a). It was suspected that these peaks were due to sequences where the drug had coupled increasingly to each of the four C bases in sequence (**132**) through the exocyclic amino function, with the most fully derivatised, most hydrophobic sequence eluting last. Peak D at t_R 17.6 min (Figure 2.4 (a)) was confirmed as conjugate (**133**) following ESMS which gave *m/z* 4842.7 consistent with the retrocycloaddition form of conjugate (**133**) (4843.8) where the drug had coupled to both the 5'-end amino function and to a single C residue through the exocyclic amino function at the C4 position. The use of 21 eq of **119** in the coupling reaction gave the best yield of the desired product. These conditions were applied to the synthesis of the temozolomide derivative (**134**) which was purified by HPLC (t_R 15.1 min) in 8% yield. ESMS analysis gave *m/z* 4846.0 which is consistent with the trisodium adduct of **134** (calc. 4846.8).

UV analysis of the isolated products (**132**) and (**134**) showed an intense absorption (λ_{max} ~ 260 nm) due to the oligonucleotides as well as much weaker absorptions with λ_{max} = 327.2 nm (**132**) and 329.1 nm (**134**) respectively, indicative of the intact drugs. Both conjugates (**132**) and (**134**) were readily purified by semi-preparative HPLC since the product peaks were distinct from all other impurities present with appreciably different retention times.

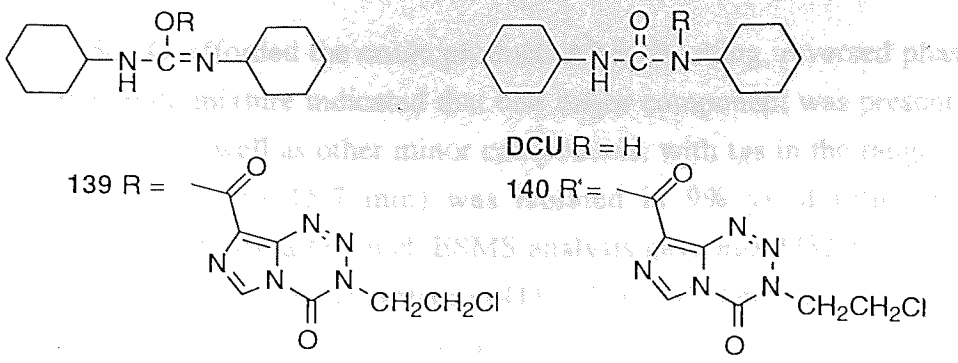
2.4 Attempted synthesis of the mitozolomide conjugate (118) directly via the phosphoramidite approach

Material (**118**) was chosen as a simple target conjugate (Scheme 2.5), to be synthesised using the drug-phosphoramidite (**135**) during the last cycle of oligomerisation. Compound (**135**) was synthesised according to Scheme 2.10.



Scheme 2.10 Synthesis of phosphoramidite (**135**). i, C_6F_5OH, DCC, 5 °C, DMF, 3 h, HOBT; ii, 5-aminopentan-1-ol, MeOH, CH_2Cl_2, r.t., 12 h; iii, 5-aminopentan-1-ol, DIPEA, PyBOP, DMF, r.t.; 50 min; iv, 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite, DIPEA, THF, r.t., 2 h.

The addition of *N,N*-dicyclohexylcarbodiimide (DCC) to a mixture of carboxylic acid (**119**), C_6F_5OH and 1-hydroxybenzotriazole (HOBT) in DMF at 5 °C afforded the pentafluorophenyl ester (**136**) after 3 h which was isolated by flash chroamtography in 63% yield (Scheme 2.10). The same procedure involving the temozolomide derivative (**120**) afforded the ester (**137**) in 23% yield. When DCC is used as a coupling agent, products are often contaminated[19] with the by-product dicyclohexyl urea (DCU). This problem is worsened by the fact that DCU is not appreciably UV active and is rather difficult to visualise on TLC plates. In this thesis work, esters (**136**) and (**137**) were isolated free of DCU contamination by using hexane-EtOAc-MeCN (6:3:1) as eluent during flash chromatography. HOBT was present in the the reactions to inhibit the formation of the unwanted DCU by-product[19] (**140**) which arises from intramolecular rearrangement of the *O*-acylisourea (**139**). HOBT reacts rapidly with **139** forming an active ester intermediate which in turn reacts with C_6F_5OH to give the desired ester (**136**) and DCU as a by-product.



Subsequent transformation of active ester (**136**) into 5-hydroxypentyl adduct (**138**) was achieved using 5-aminopentan-1-ol as the nucleophile but in very modest yield (30%). A higher yielding synthesis of **138** was provided by reacting carboxylic acid (**119**) directly with 5-aminopentan-1-ol using PyBOP as coupling agent in the presence of DIPEA. Hydroxypentyl adduct (**138**) was isolated in 60% yield after purification by flash chromatography. Phosphitylation at the primary hydroxyl group of **138**, using 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite, afforded phosphoramidite (**135**) in 62% yield (35% overall yield over three steps from **119**) for use directly in automated solid phase synthesis of the 5'-end conjugate sequence (**118**).

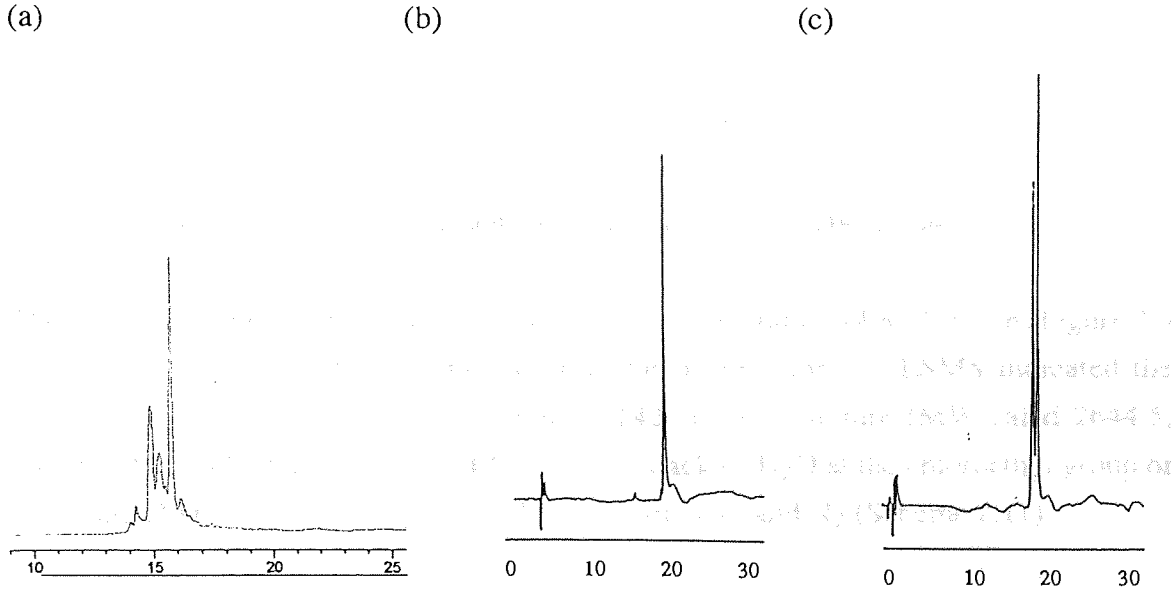
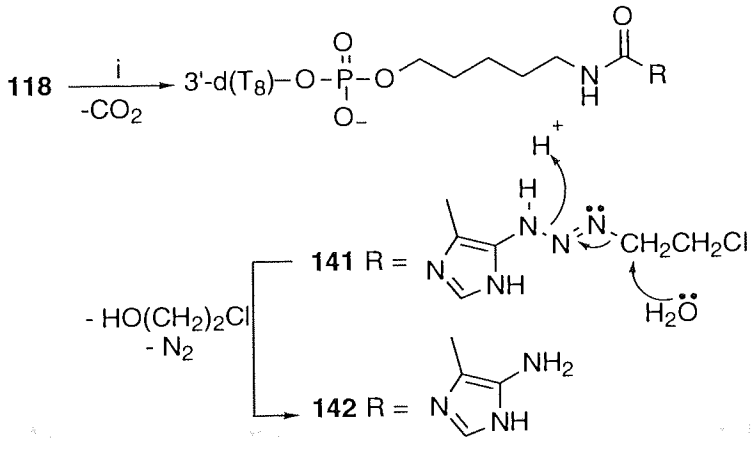


Figure 2.5 HPLC profiles of (a) crude oligonucleotide (**141**); (b) oligonucleotide (**141**) following purification and (c) co-injection of (b) and d(T_8).

Conventional solid phase synthesis of **118** was attempted using phosphoramidite (**135**) on SLCPG support (**95**). Fluoride-catalysed cleavage and deprotection using TBAF

during 2 h at 50 °C, afforded the crude product. After desalting, reversed phase HPLC analysis of the crude mixture indicated that one major component was present (t_R 15.7 min, Figure 2.5 (a)) as well as other minor components, with t_Rs in the range 14.8-15.3 min. The major peak (t_R 15.7 min) was isolated in 9% yield following HPLC purification (Figure 2.5 (b)) afterwhich ESMS analysis gave *m/z* 2732.5 consistent with that calculated for the triazene derivative (**141**) (2734.5) (Scheme 2.11) which formed presumably as a result of nucleophilic attack by H_2O at the C4 carbonyl of the imidazotetrazinone ring of **118**, followed by decarboxylation (Scheme 1.1). Co-injection of sequence (**141**) with an authentic sample of the unconjugated parent sequence d(T_8) showed two major peaks (t_Rs 18.0, 18.5 min) (Figure 2.5 (c))[119], the more hydrophobic conjugated sequence (**141**) having the longer retention time as expected.



Scheme 2.11 Decomposition of the mitozolomide conjugate (**118**) in water. i, H_2O.

The remaining peaks of weaker intensity, with t_Rs in the range 14.8-15.3 min (Figure 2.5 (a)), were also collected and combined together in one portion. ESMS indicated the presence of the aminoimidazole conjugate (**142**) in this mixture (MW calcd 2644.5; found 2644.0) which formed from **141** through attack of H_2O at the chloroethyl group of conjugate (**141**) with subsequent loss of 2-chloroethanol and N_2 (Scheme 2.11).

It was therefore clear that the fluoride cleavage/deprotection conditions were not compatible with the drug. In order to determine that coupling of phosphoramidite **135** at the 5'-end during DNA synthesis was successful, a portion of the fully protected sequence was cleaved from the SLCPG support using conditions under which the tetrazinone ring is stable (2% TFA-CH_2Cl_2, 1 h). The solvent was evaporated and the residue analysed by HPLC. One main peak was observed at t_R 27.0 min, the long t_R consistent with the presence of hydrophobic cyanoethyl protecting groups on

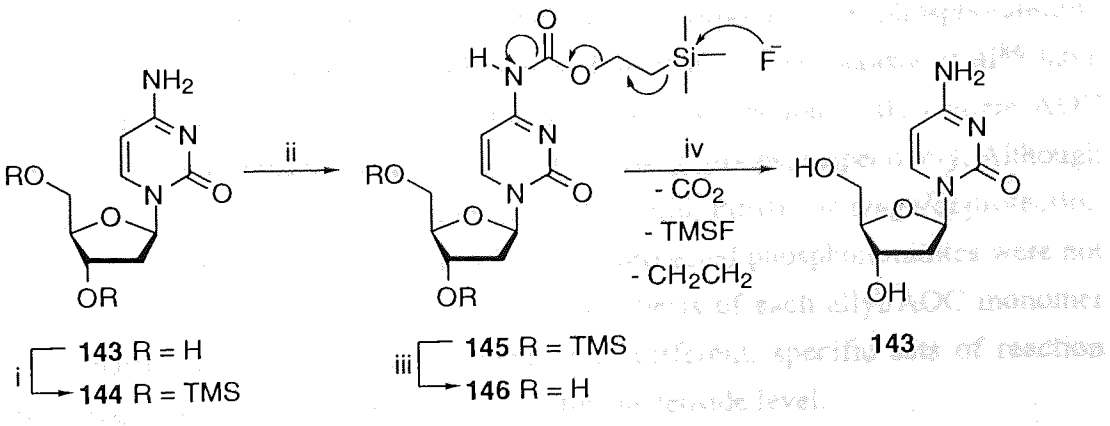
phosphorus. UV analysis of this sequence (200-400 nm) showed 2 maxima (λ_{max} = 264.8, λ_{max} = 323.3 nm). The intense maximum at 264.8 nm was consistent with the oligonucleotide while the weaker absorpotion at 323.3 nm was consistent with the characteristic band in the UV spectrum of the free drug (**1**) (λ_{max} = 325 nm).[7] A portion of this sample, when treated to conc. NH_3 (aq) solution at 55 °C during 1 h, showed a single peak (t_R 15.0 min) in the HPLC profile and a single absorption maximum (λ_{max} = 268 nm) in the UV spectrum with no absorbance between 300-350 nm. This was consistent with the known base lability of the imidazotetrazinones, where loss of the 300-350 nm absortion band accompanies ring-opening of the tetrazinone ring. Therefore, incorporation of the drug via the phosphoramidite method was successful, but the fluoride cleavage/deprotection conditions were too severe for the base-labile drug.

2.5 Fluoride-labile, DNA-base protecting group

2.5.1 The trimethylsilylethoxycarbonyl (TEOC) protecting group

In parallel with this work, the problem of how to incorproate DNA bases (A, G, C) other than T into a drug-conjugate sequence was addressed. The exocyclic amino protecting groups of stanadard DNA monomers (Bz, Bu^i, Ac) may be deprotected using conc. NH_3 (aq) but not TBAF-AcOH. Thus an SLCPG-compatible base protecting group was sought which would allow use of the fluoride cleavege/deprotection protocol. The trimethylsilylethoxycarbonyl[120, 121] (TEOC) protecting group has been successfully cleaved using TBAF and so was considered a suitable DNA base protecting group.

2.5.2 Synthesis of TEOC-dC (146)



Scheme 2.12 Synthesis and fluoride-catalysed deprotection of **146**. i, TMSCl, pyridine, 0 °C, 90 min; ii, 2-(trimethylsilyl)ethyl *p*-nitrophenyl carbonate, 60 °C, 12 h; iii, H_2O, MeOH, r.t., 3 h; iv, TBAF (0.5 M)-AcOH (0.5 M), THF, 55 °C, 2 h.

2'-Deoxycytidine (**143**) was protected with TEOC as shown in Scheme 2.12. Firstly, the 3'-and 5'-hydroxyl groups of **143** were transiently protected with TMS by reaction with TMSCl at 0 °C during 90 min to give **144**. The 4-amino group of **144** was protected with TEOC by reaction with 2-(trimethylsilyl)ethyl *p*-nitrophenyl carbonate at 60 °C during 12 h affording the protected nucleoside (**145**). Hydrolysis of the TMS protecting groups using H_2O and MeOH followed by flash chromatographic purification, afforded the desired product **146** in 65% yield.

2.5.3 Removal of the TEOC protecting group from TEOC-dC (146)

The suitability of TEOC as an SLCPG/DNA-compatible base protecting group was assessed by treating **146** to the standard TBAF-AcOH cleavage/deprotection conditions. A quantity of the nucleoside (**146**) was dissolved in the deprotection solution and the reaction was monitored by TLC. After 1 h 10 min at ambient temperature, only the starting material was observed by TLC. The mixture was then heated for 2 h at 55 °C, after which, TLC indicated the presence of a single, new spot (R_f 0.07) which was shown to be 2'-deoxycytidine (**143**) by comparison with an authentic sample of this nucleoside.

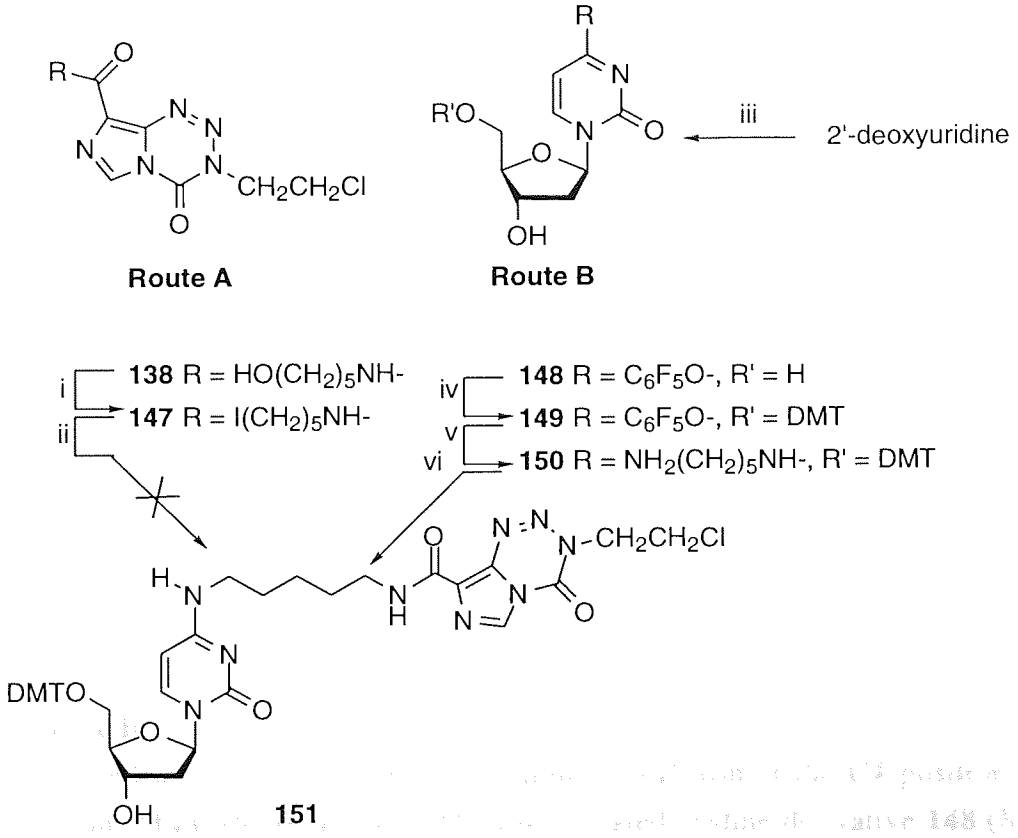
Attack by fluoride at the Si centre led to formation of the volatile by-products, trimethylsilylfluoride (TMSF), ethene and CO_2 as well as the free nucleoside (**143**). Thus, TEOC was considered a suitable base protecting group for 2'-deoxycytidine. However, as the drug (**1**) was found to be unstable towards the same fluoride deprotection conditions, the incorporation of **1** into sequences directly, using phosphoramidite methodology, was abandoned.

Other methods could have been used for the direct incorporation of phosphoramidite derivatives of the drugs into oligonucleotides. For example, Hayakawa et al[86] have synthesised oligonucleotides using monomers of type **84** (Section 1.10.2) where AOC and allyl serve as nucleobase and phosphorus protecting groups, respectively. Although the drugs may have been compatible with the mild Pd(0) cleavage/deprotection conditions used by Hayakawa et al, the allyl/AOC protected phosphoramidites were not commercially available. The necessity for the synthesis of each allyl/AOC monomer made this particular approach inconvenient with different, specific sets of reaction conditions required for different DNA bases at the nucleoside level.

2.6 Synthetic routes towards a drug-nucleoside phosphoramidite

2.6.1 Coupling of mitozolomide to dC - attempted synthesis of 151

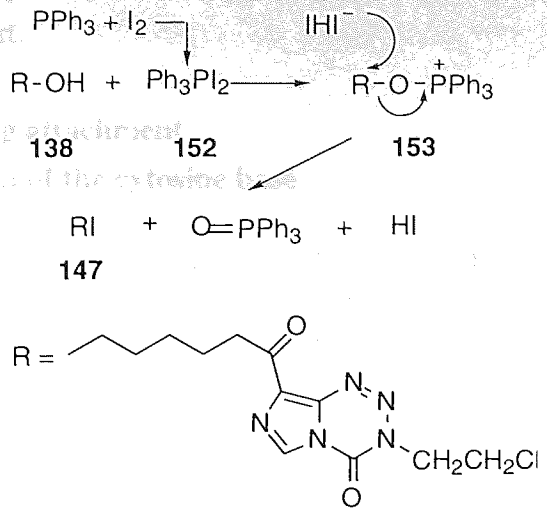
The following work was carried out in parallel with the above work described in Section 2.4. In order to attach the drugs at the C4 position of 2'-deoxycytidine, two different routes (Scheme 2.13, routes A and B) were followed with the overall aim of synthesisng a nucleoside-drug phosphoramidite.



Scheme 2.13 Synthetic routes towards nucleoside conjugate (**151**). i, PPh_3, MeCN, DIPEA, I_2, r.t., 12 h; ii, 2'-deoxycytosine (**143**), TMSCl, pyridine, 50 °C, 24 h; iii, $CF_3C(O)_2O$, pyridine, C_6F_5OH, 135 h; iv, DMTCl, pyridine, r.t., 1.5 h, MeOH; v, $H_2N(CH_2)_5NH_2$, 1,4-dioxane, 12 h, 70 °C; vi, **136**, DMSO, r.t., 1 h.

2.6.2 Route A

Since the hydroxy pentyl derivative (**138**) was readily available, it was converted to the iodo derivative (**147**) in 31% yield by reaction with iodine in the presence of PPh_3 following a literature procedure (Scheme 2.13).[122] The reaction mechanism[123, 124] (Scheme 2.14) involved reaction of PPh_3 with I_2 to form PPh_3I_2 (**152**) which following subsequent reaction with **138** afforded the phosphonium ion (**153**). Attack by iodide at the electrophilic R group of **153** afforded the highly stable PPh_3=O and the desired iodoalkyl derivative (**147**).



Scheme 2.14 Mechanism of formation of **147**.

An attempt was then made to form the desired nucleoside-drug conjugate (**151**) by mixing **147** with 2'-deoxycytosine (**143**) in the presence of pyridine and TMSCl at 50 °C, during 24 h. However, TLC analysis indicated that no reaction had taken place after this time. Conjugation of the lone electron pair of the exocyclic amino group of **143** with the π–deficient pyrimidine ring system results in a weakly nucleophilic N atom which could not displace the iodo group of **147** under the conditions used.

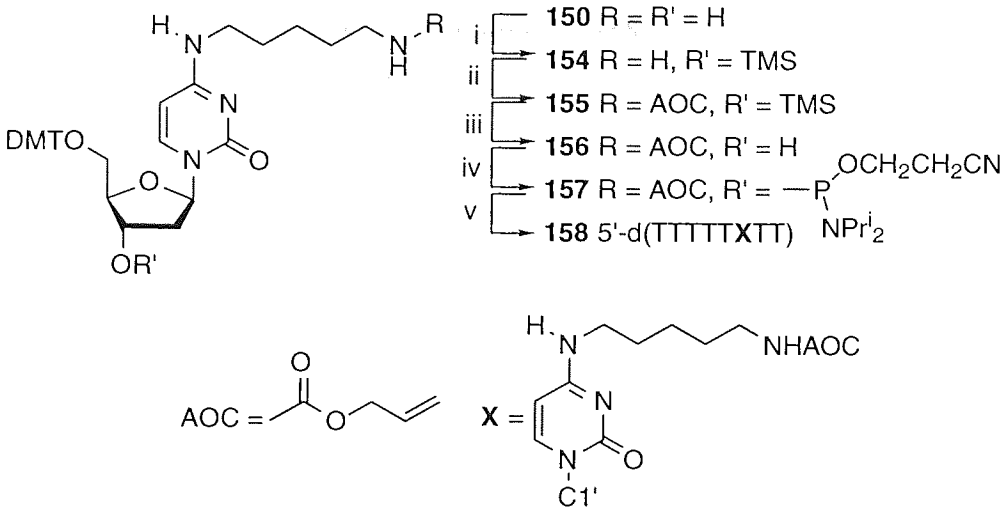
2.6.3 Route B

It was decided to insert a more reactive, amino alkyl arm at the C4 position of 2'-deoxycytosine (**143**) (Scheme 2.13). The C4 activated uridine derivative **148** (Scheme 2.13) is known to provide cytidine derivatives through displacement of the pentafluorophenoxy group by amine nucleophiles.[59, 125] **148** was synthesised in 72% yield from 2'-deoxyuridine according to a literature procedure,[125] which, following reaction with 1,5-diaminopentane at 70 °C during 12 h, afforded a highly polar product which was difficult to purify. **148** was thus elaborated to **149** in 64% yield according to a literature procedure,[59] which, following reaction with 1,5-diaminopentane at 70 °C during 12 h, afforded the desired product (**150**) in 85% yield following short-column flash chromatographic purification. **150** was then reacted with mitozolomide ester (**136**) during 1 h, after which the product was purified by flash chromtography affording the desired conjugate (**151**) in 31% yield as judged by ^{1}H NMR. Although phosphitylation of the 3'-OH function was all that was required to convert **151** to the desired DNA building block, other work had shown that the TBAF-AcOH cleavage/deprotection conditions were incompatible with the drug (Section 2.4). So this area of work was

abandoned in favour of post-synthetic drug conjugation to a fully deprotected sequence bound to SLCPG support.

2.7 Post-synthetic drug attachment

2.7.1 Functionalisation of the cytosine base



Scheme 2.15 Synthesis of phosphoramidite (**157**) and sequence (**158**). i, TMSCl, pyridine, 0 °C, 40 min; ii, allyloxy carbonyl chloroformate, r.t., 17 h; iii, H_2O, MeOH, r.t., 3 h; iv, 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite, DIPEA, THF, r.t., 105 min; v, DNA synthesis (succinyl CPG), conc. NH_3 (aq), 55 °C, 16 h.

It was necessary to suitably protect the free amino function of **150** such that the modified nucleoside could be incorporated into a sequence using phosphoramidite methodology and deprotected while still bound to SLCPG support. Both acid (DMT, MMT) and base-labile (trifluoroacetyl) protecting groups were considered unsuitable as the necessary deprotection conditions for each would either have caused support-cleavage (base, strong acid) or interference with DNA synthesis (acid).

2.7.2 The allyloxycarbonyl (AOC) protecting group

The AOC group may be cleaved under mild conditions using Pd(0),[86] which made it suitable for this work. The 3'-OH function of nucleoside (**150**) was transiently protected by reaction with TMSCl at 0 °C affording **154** (Scheme 2.15). The free amino function was then protected by reaction with allyloxycarbonylchloroformate during 17 h, forming the protected carbamate (**155**). Following hydrolysis of the TMS group and chromatographic purification, the desired product (**156**) was isolated in 46% yield. Phosphitylation at the secondary hydroxyl group of **156**, using 2-cyanoethyl *N,N*-

diisopropylchlorophosphoramidite, afforded the phosphoramidite (**157**) in 39% yield following chromatographic purification. The product was 69% pure based on ^{31}P NMR analysis.

157 was used directly in the synthesis of the model sequence (**158**) on succinyl CPG support. The sequence was cleaved and deprotected using conc. NH_3 (aq) at 55 °C during 16 h affording the allyl-protected product (**158**) (t_R 15.0 min) in 81% crude yield. Product identity was confirmed by ESMS analysis (MW found 2525.2; calcd 2524.5). Due to a lack of time, sequence (**158**) was not synthesised on SLCPG support and so the compatibility of the allyloxycarbonyl deprotection conditions with this support could not be assessed further.

2.8 Other routes towards nucleoside-mitozolomide conjugates for sequence synthesis via the phosphoramidite approach

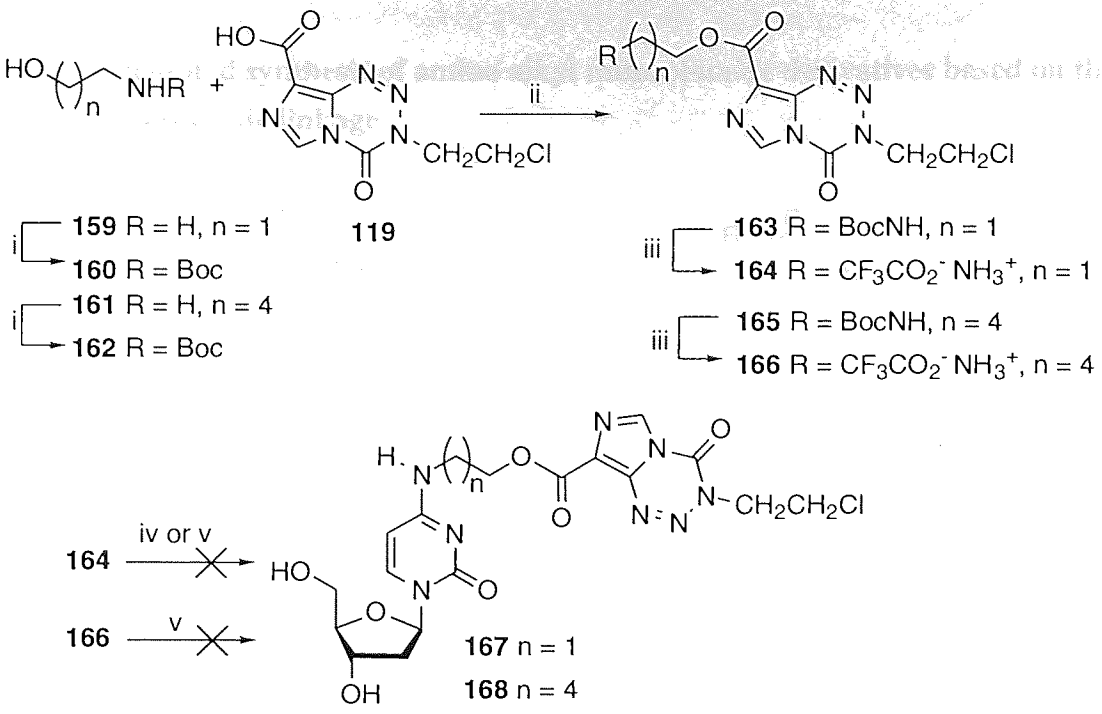
2.8.1 Conjugates based on the carboxylate linkage

Other routes towards the synthesis of nucleoside-drug conjugates were followed which centred around the synthesis of drug derivatives containing an amino-alkyl linker arm which was to be coupled to the activated uridine derivative (**148**).

Although it has been suggested[6] that the C8 amido function of **1** is important for DNA binding of the free drug (Section 1.2.4), it was considered that this function was not a necessity for DNA targeting of the oligonucleotide-drug conjugate since the oligonucleotide was to provide site-specific delivery of the drug to the target sequence. As a result, attempts were made to synthesise conjugates (**167**) and (**168**) (Scheme 2.16) in which an ester function replaces the more familiar carboxamide group at C8 position of the drug.

2.8.2 Boc as an amino protecting group

The *t*-butoxycarbonyl (Boc) group has been used successfully by Clark et al[16] to insert *N*1,*N*8-bis(*t*-butoxycarbonyl)-*N*4-(3-aminopropyl)spermidine at the C8 position of **1** and **2**. The Boc protecting groups were subsequently cleaved using TFA, demonstrating the compatibility of this protecting group with the drugs. 2-Aminoethan-1-ol and 5-aminopentan-1-ol were reacted with 2-(*t*-butoxycarbonyloxyimino)-2-phenylacetonitrile (Boc-ON) affording the *N*-Boc-protected alcohols (**160**) and (**162**), in yields of 75% and 79% respectively. PyBOP was used to couple the alcohols (**160**) and (**162**) to the 8-acid



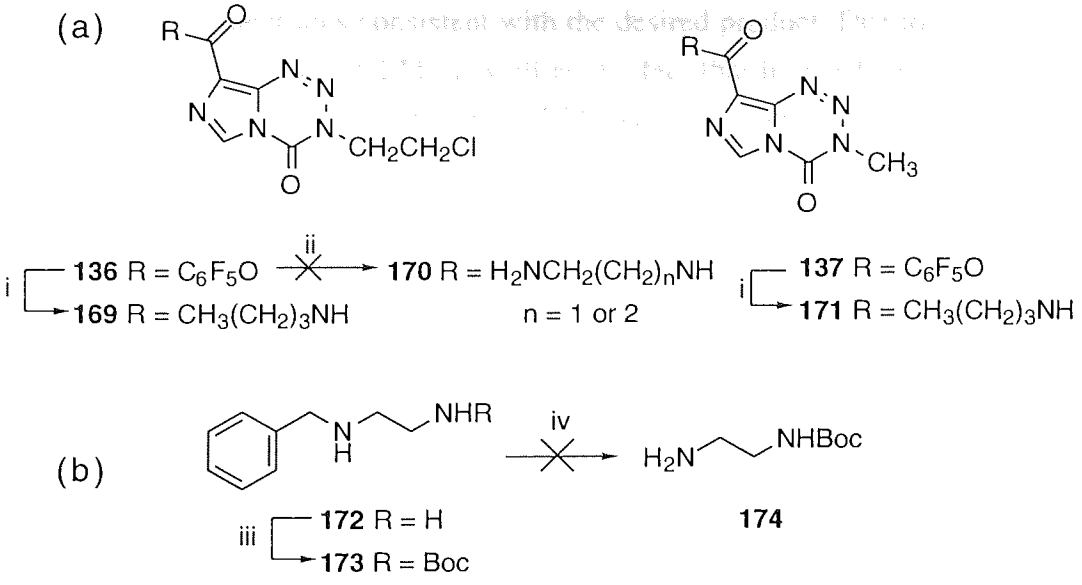
Scheme 2.16 Attempted syntheses of mitozolomide conjugates (**167**) and (**168**). i, Boc-ON, 1,4-dioxane, r.t., 3 h; ii, PyBOP, DIPEA, DMF, r.t., 45 min; iii, TFA-CH_2Cl_2 (9:1), 1 h; iv, **148**, Et_3N, DMAP, 1,4-dioxane, 70 °C, 15 h; v, **148**, DIPEA, 1,4-dioxane, 70 °C, 12 h.

derivative of mitozolomide (**119**) affording **163** and **165** respectively, with a yield of 70% for the latter. The Boc-protecting groups on both **163** and **165** were cleaved using TFA-CH_2Cl_2 (9:1) during 1 h. The reaction mixtures were partitioned between H_2O and CH_2Cl_2. The desired products (**164**) and (**166**) were isolated, following evaporation of the aqueous phase, as the trifluoroacetyl ammonium salts in satisfactory yield (73% for **164** over 2 steps, and 99% for **166** over 1 step, 72% over 2 steps).

2.8.3 Attempted synthesis of the mitozolomide conjugates (167) and (168) of 2'-deoxycytosine

Attempts were made to couple **164** to the activated nucleoside (**148**) in the presence of Et_3N (3 eq) at 70 °C, during 3 h after which time TLC analysis only indicated the presence of starting materials. DMAP (1.6 eq) was then added to the reaction mixture which was continued for a further 12 h. TLC analysis after this time did not indicate the presence of product. This reaction was repeated using DIPEA (3 eq) as a base in place of Et_3N and DMAP but failed to provide the desired product. The reaction was repeated using (**166**) and (**148**) in the presence of DIPEA (5 eq) at 70 °C during 12 h, but again the reaction failed. This route was not continued beyond this stage.

2.8.4 Attempted synthesis of amino alkyl mitozolomide derivatives based on the carboxamide linkage



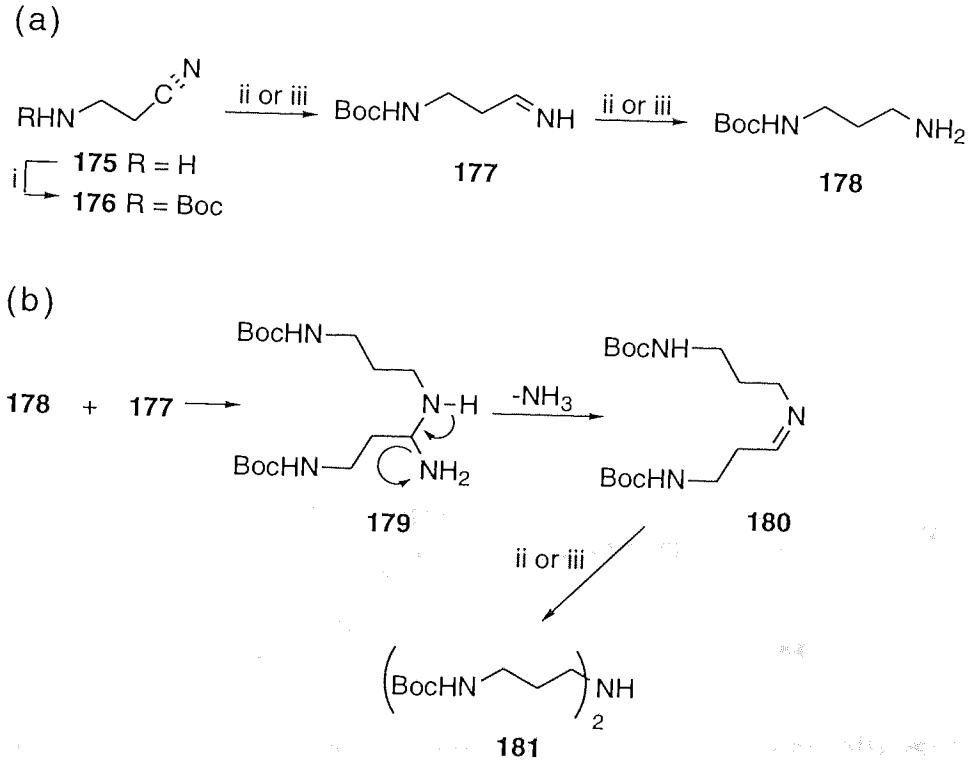
Scheme 2.17 (a) Reactions of the active esters (**136**) and (**137**) with amine nucleophiles and (b) attempted synthesis of **170**. i, CH_2Cl_2, $BuNH_2$, r.t, 45 min; ii, CH_2Cl_2, $H_2NCH_2(CH_2)_nNH_2$ (n = 1 or 2), r.t., 45 min; iii, Boc-ON, 1,4-dioxane, r.t., 3 h; iv, MeOH, 10% Pd on C catalyst, r.t., H_2, 7 days.

The reactivity of the pentafluorophenoxy esters (**136**) and (**137**) towards amine nucleophiles was assessed by reaction with Bu^nNH_2 during 45 min after which time the desired products (**169**) and (**171**) were isolated in satisfactory yields of 60% and 58% respectively. However, dropwise addition of a solution of ester (**136**) (1 eq) in CH_2Cl_2 to an excess of either ethylenediamine (**170**) (n = 1) (16 eq) or 1,3-diaminopropane (**170**) (n = 2) (16 eq) during 45 min resulted in degradation through ring-opening of the tetrazinone ring as evidenced by the lack of the characteristic[12] C6-H singlet at ~ 8.9 ppm in the ^1H NMR spectrum. Diamines (**170**) (n = 1, 2) were used in excess to prevent dimer formation but this excess also led to degradation by attack of the amine nucleophiles at the C4 carbonyl of the tetrazinone ring (see Scheme 1.2 (b)).[7,10]

In order to avoid this, synthesis of mono Boc-protected diamines such as **174** was attempted (Scheme 2.17 (b)), where the protecting group could be cleaved from the resulting drug derivative. Towards **174**, reaction of benzylethylenedianine (**172**) with Boc-ON afforded *t*-butyl 2-benzylaminoethylamine-*N*-carboxylate (**173**) in 74% yield, which, following hydrogenolysis using 10% Pd on C during 7 days afforded a viscous

yellow oil in 75% yield. The ^{1}H NMR spectrum of the product was identical to that of an authentic sample of **174**, but curiously, a doublet (0.94 ppm) and a septet (2.65 ppm) were also prominent in the spectrum. Irradiation of the septet resulted in collapse of the doublet to a singlet. Also, the mass spectral analysis of the product did not yield a molecular ion or fragment ions consistent with the desired product. Due to the problems associated with the synthesis of **174**, as well as the fact that higher homologues of the benzyl-protected diamino starting material (**172**) were not commercially available, an alternative route to the Boc-protected propyl homologue was sought.

2.8.5 Attempted synthesis of the Boc-protected amine (178)

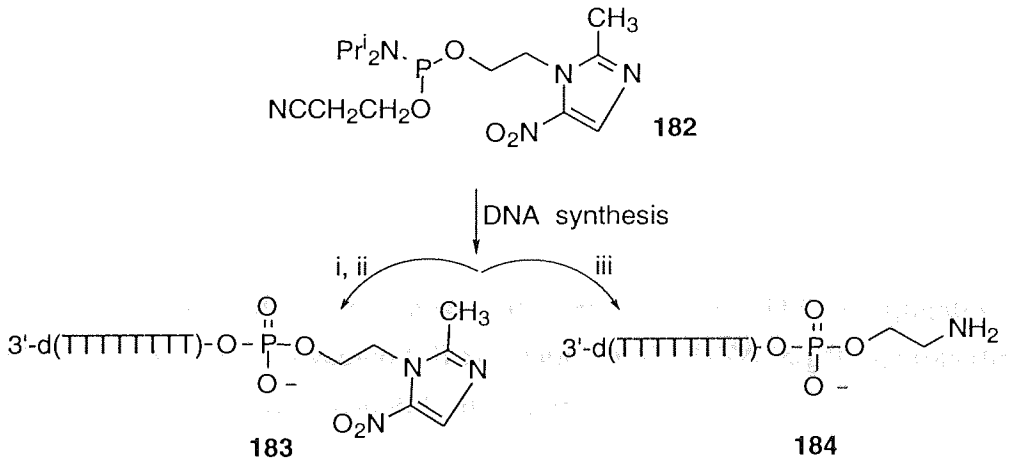


Scheme 2.18 (a) Attempted synthesis of **178** and (b) mechanism of formation of the secondary amine (**181**). i, Boc-ON, 1,4-dioxane, 5 °C, 3 h; ii, MeOH, 10% Pd on C, H_2, r.t., 6 days; iii, 10% NH_3-MeOH, 5% Rh on C, H_2, 80 psi, 60 h.

3-Aminopropionitrile (**175**) was reacted with Boc-ON affording the protected nitrile (**176**) in 67% yield. Attempted reduction of the nitrile group of **176** to give **178** in the presence of 10% Pd on C under a hydrogen atmosphere during 6 days afforded a complex mixture containing 5 components evidenced by TLC analysis. No attempt was made to purify the product mixture. A more complex product mixture, 8 components by TLC analysis, was obtained by carrying out the reaction in 10% NH_3-MeOH under a

hydrogen atmosphere, in the presence of 5% Rh on C, at 80 psi during 60 h. A number of side-reactions were possible during reduction of the nitrile group, the most common being the formation of secondary amines, although tertiary amine formation was also possible.[126] Secondary amine formation possibly occurred by reaction of the desired primary amine (**178**) with the imine (**177**) from the partially reduced nitrile to give **179** (Scheme 2.1 (b)). Subseqeunt loss of NH_3 afforded **180** which, following hydrogenolysis of the imine function afforded the secondary amine (**181**) (Scheme 2.18). Tertiary amine formation would also be possible by reaction of the secondary amine (**181**) with the imine (**177**) by a mechanism similar to that for secondary amine formation. Some of the many spots observed by TLC analysis probably correspond to the species (**177-181**). Difficulty with this reduction step prompted the search for the alternative route (Section 2.6.3) which was successfully used to synthesise conjugate (**151**) (Scheme 2.13).

2.9 Synthesis of the metronidazole conjugate (183)



Scheme 2.19 Synthesis of sequences (**183**) and (**184**). i, 20% Et_3N-pyridine; ii, conc. NH_3 (aq), r.t., 30 min; iii, conc. NH_3 (aq), 60 °C, 16 h.

As well as its proven antibacterial activities, metronidazole[127] is an agent which is clinically active against hypoxic tumours and interacts directly with DNA and it was therefore of interest to attach this agent to TFO sequences to assess any change in reactivity compared with the parent drug. The metronidazole phosphoramidite (**182**) was used to synthesise the model sequence (**183**) on succinyl CPG support. A portion of the protected, support-bound sequence was treated with 20% Et_3N-pyridine during 2 h to cleave the cyanoethyl groups. Conc. NH_3 (aq) was then used to cleave the fully deprotected sequence from the support at r.t. during 30 min. Following semi-preparative HPLC purification, the resulting product (t_R 13.3 min) (Figure 2.6) was analysed by ESMS which gave *m/z* 2604.2 consistent with the target conjugate (**183**) (calc. 2603.4).

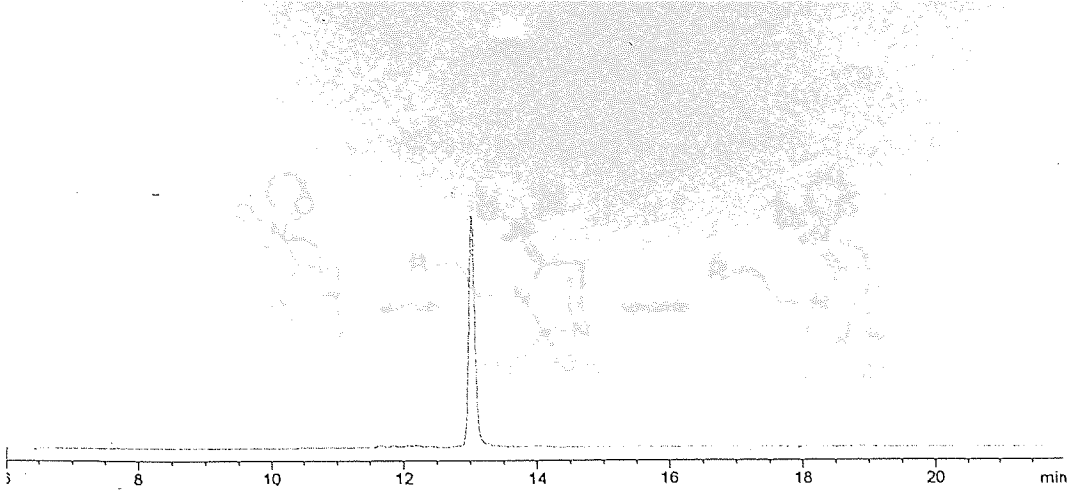


Figure 2.6 HPLC profile of the metronidazole conjugate (**183**).

Another portion of the protected, support-bound sequence was treated with conc. NH_3 (aq) at 60 °C during 16 h. The purified product (t_R 12.2 min) was analysed by ESMS which confirmed its identity as **184** (MW found 2494.1; calcd 2492.4). A possible mechanism for the formation of **184** from **183** is shown in Scheme 2.20. As a result of the electron-withdrawing, activating nature of the nitro group, addition of NH_3 at C1 of the 5-membered ring would lead to formation of the intermediate (**185**). Ring-opening at the C1-N5 bond would lead to the formation of **186** which following cyclisation affords the intermediate (**187**). Following proton abstraction by NH_3 from **188** the product (**184**) is formed as well as the substituted imidazole (**189**).

Although a direct method for synthesis of metronidazole-TFO conjugates was established, time constraints prevented full evaluation of the DNA targeting properties of **183** and synthesis of other oligonucleotide analogues.

2.10 Evaluation of TFO conjugates (132) and (134)

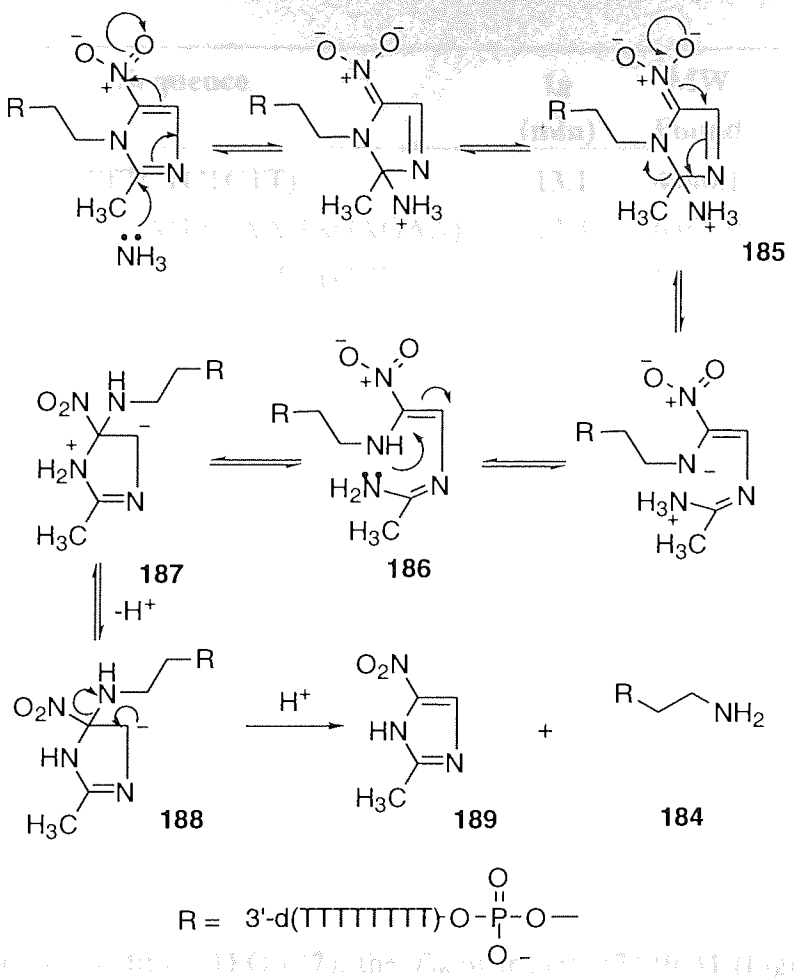
2.10.1 Objectives

Having successfully synthesised and characterised the drug conjugates (**132**) and (**134**), it remained to evaluate them:

1. To assess whether the conjugate sequences could form stable triple helcies with a suitable duplex DNA target.

2. To incubate the conjugates with the target duplex and thereafter determine whether any modification of the duplex target had occurred using ESMS analysis.

3. To compare the reactivity of the conjugates with that of the free drugs (**1**) and (**2**).



Scheme 2.20 Mechanism for the formation of the oligonucleotide (**184**) following prolonged treatment with conc. NH_3 (aq).

2.10.2 Targeting of drug-conjugates (132) and (134) to duplex DNA

5'-d(TTTCTTTTCTCTCTT) **27**

5'-d(GGG**G**GAAAGAAAAGAGAGAA) **30**

3'-d(CCCCCTTTCTTTTCTCTCTT) **31**

The drug conjugates (**132**) and (**134**) were evaluated using the duplex **30.31** with the aim of alkylating the G base indicated. Sequences (**27**, **30** and **31**) were synthesised on normal CPG support and cleaved and deprotected using conc. NH_3 (aq) prior to semi-preparative HPLC purification and characterisation by ESMS (Table 2.2).

Table 2.2 HPLC and ESMS analysis of oligonucleoides (**27**), (**30**) and (**31**).

	Sequence	t_R (min)	MW Found	MW Calcd
27	5'-d(TTTCTTTTCTCTCTT)	13.1	4440.1	4438.7
30	5'-d(GGGGGAAAGAAAAGAGAGAA)	12.3	6369.0	6365.2
31	3'-d(CCCCCTTTCTTTTCTCTCTT)	13.3	5885.7	5884.0

Thermal denaturating experiments (Table 2.3) were carried out using PIPES buffer (pH 6.4) in the presence of spermine, Mg^{2+} and Na^{+} using the sequences listed in Table 2.3 (1.5 μM) each .

Table 2.3 Melting transitions recorded for triplexes **27*30.31**, **132*30.31** and **134*30.31**.

Triplex	$T_{m\,1}$ (°C)	$T_{m\,2}$ (°C)
27*30.31	46.7 ± 0.5	65.9 ± 0.5
132*30.31	48.4 ± 0.5	65.7 ± 0.5
134*30.31	47.0 ± 0.5	65.8 ± 0.5

In the case of the unmodified TFO (**27**), the T_m of triplex **27*30.31** (Figure 2.7) was found to be 46.7 ± 0.5 °C. For all triplexes studied, the duplex T_m ~ 65.7 ± 0.5 °C was unaffected by the choice of TFO. Typical melting profiles obtained when the TFO drug-conjugates (**132**) and (**134**) were incubated with the duplex **30.31** can be seen in Figure 2.7. The triplex T_ms were determined to be 48.4 ± 0.5 °C for the triplex containing **132** and 47.0 ± 0.5 °C for the triplex containing **134**. The small difference between the triplex melting temperatures of the conjugates (< 1.7 °C) and that of the unmodified sequence indicates that triplex stability was not compromised by the presence of the drugs at the 5'-end. It also suggests that no unfavourable interaction, such as intercalation, of drugs with the duplex target occurs. Thus, it remained to assess whether the drugs could alkylate the middle G of the GGG target in sequence (**30**) of the duplex strand.

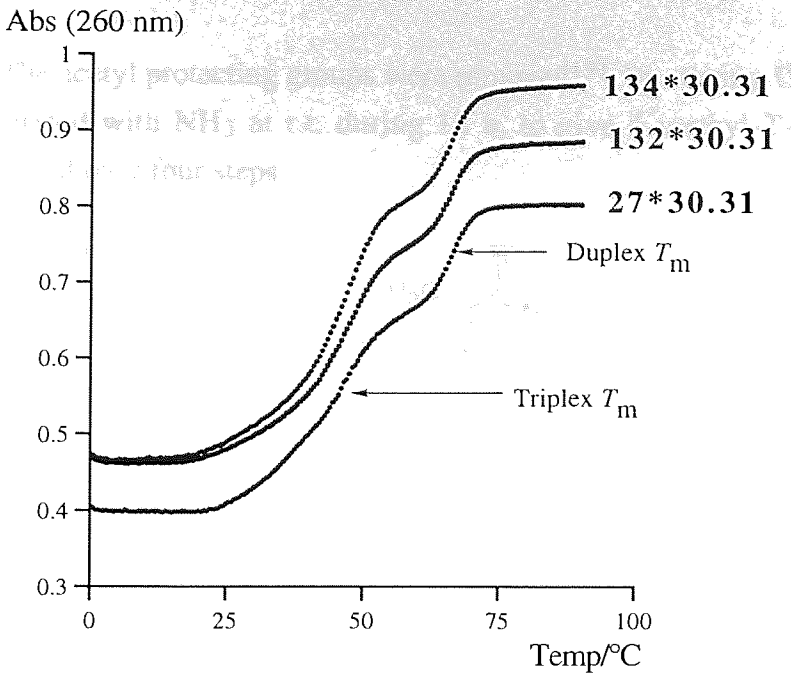


Figure 2.7 Melting profiles of triplexes **27*30.31**, **132*30.31** and **134*30.31**.

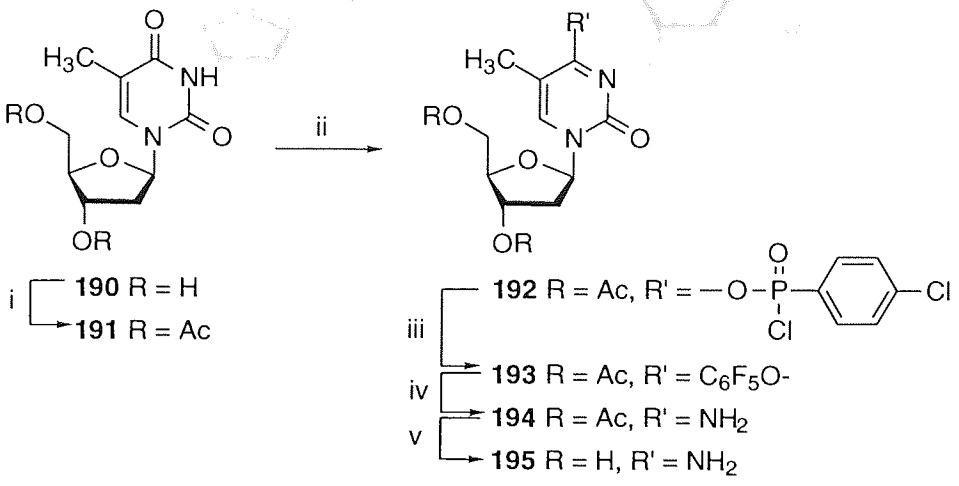
The conjugates **132** and **134** as well as the free drugs (**1**) and (**2**) were incubated seperately with the duplex **30.31** at 24 °C during 3 days. The buffer conditions were identical to those used above, except that the parent drugs (**1**) and (**2**) and sequences were each present at a concentration of 10 μM. After this time, all UV active material present in the samples was isolated in one portion by semi-preparative HPLC and desalted using a NAP-10 column. Analysis of the mixtures by ESMS failed, possibly due to the low concentrations of the samples. It was anticipated that ESMS analysis at the 5-10 μM level would be possible since others[75] had obtained satisfactory results using ESMS at similar concentration levels. More rigourous desalting procedures were perhaps necessary for ESMS analysis at this concentration level.

2.11 Towards drug-conjugate targeting under physiological conditions

2.11.1 Synthesis of 5-methyl-2'-deoxycytosine (195)

So that the drug conjugates could be analysed at physiological pH, $5^{m}C$ was chosen as a replacement for C in TFO (**27**). Towards this, 5-methyl-2'-deoxycytosine (**195**) was synthesised (Scheme 2.21) using 2'-deoxythymidine (**190**) as starting material. Reaction of **190** with acetic anhydride during 18 h at 0 °C afforded the 3', 5'-diacetyl derivative (**191**)[128,129] which was further reacted with 4-chlorophenyl phosphorodichloridate[130] during 48 h to give **192**, which was not isolated. Addition of C_6F_5OH to the reaction mixture afforded the pentafluorophenyl derivative (**193**)[130] after 7 days which was isolated and then stirred in a mixture of 1,4-dioxane and conc. NH_3 (aq) (3:1) in a sealed bomb at r.t during 4 h which resulted in displacement of the C_6F_5O group by NH_3 to

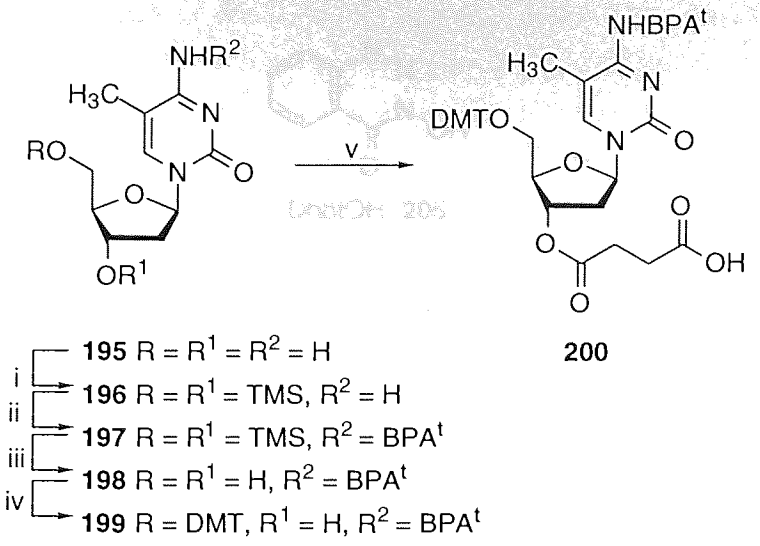
form **194**.[131] The acetyl protecting groups were removed[131] by stirring **194** in a solution of MeOH saturated with NH_3 at r.t. during 12 h, to give 5-methyl-2'-deoxycytosine (**195**), in 51% yield over four steps.



Scheme 2.21 Synthesis of nucleoside (**195**). i, pyridine, Ac_2O, 0 °C, 18 h; ii, pyridine, 4-chlorophenyl phosphorodichloridate, 0 °C, 48 h; iii, C_6F_5OH, r.t., 7 days; iv, 1,4-dioxane-conc. NH_3 (aq) (3:1), r.t., 4 h; v, MeOH-NH_3 (sat.), r.t., 12 h.

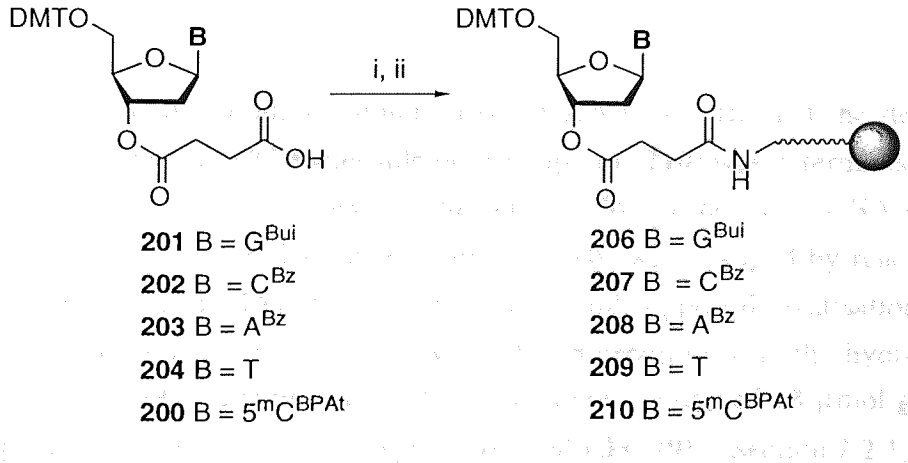
2.11.2 Synthesis of the protected nucleoside (199) and the 3'-*O*-succinate (200)

For DNA/conjugate synthesis, the use of the BPA^t group for protection of the 4-amino function of **195** was necessary. The 3'- and 5'-OH groups of **195** were TMS-protected by reaction with TMSCl at 0 °C during 1 h to give **196** (Scheme 2.22). Following addition of BPA^t anhydride to the reaction mixture at 0 °C, TLC analysis indicated that the BPA^t-protected nucleoside (**197**) had formed after 3.5 h. The TMS groups were hydrolysed and the product mixture was concentrated. Reaction of the crude product (**198**) with DMTCl during 24 h afforded the protected nucleoside (**199**) which was isolated by flash chromatography, eluting with EtOAc-EtOH-Et_3N (80:20:1) in 25% yield over 2 steps. The 3'-OH function of **199** was derivatised by reaction with succinic anhydride in the presence of DMAP to give the succinate ester (**200**) which was used directly to derivatise LCAA-CPG.



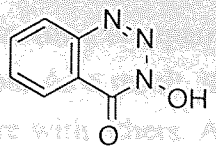
Scheme 2.22 Synthesis of the protected nucleoside (**199**) and the 3'-*O*-succinate (**200**). i, TMSCl, pyridine, 0 °C, 1 h; ii, *t*-butylphenoxyacetic anhydride, 0 °C, 3 h 30 min; iii, MeOH, H_2O, r.t., 20 min; iv, DMTCl, pyridine, r.t., 24 h; v, DMAP, succinic anhydride, pyridine, r.t., 48 h.

2.11.3 Derivatisation of LCAA-CPG using the 3'-*O*-succinates (200-204)



Scheme 2.23 Solid support derivatisation using 3'-*O*-succinates (**200-204**).
i, DMAP, DhbtOH, DCC, DMF, LCAA-CPG, r.t., 24 h; ii, Ac_2O, DMAP, pyridine, r.t., 1 h.

The 3'-*O*-succinate (**200**) nucleoside was attached to CPG support by a novel method developed during the course of this work. It was found that the commercially available 3'-*O*-succinate derivatives (**201-204**) (Scheme 2.23) could be coupled with high efficiency, during 24 h, in the presence of 3,4-dihydro-3-hydroxy-4-oxo-1,2,3-benzotriazine (DhbtOH, **205**) to LCAA-CPG to give the derivatised supports with high loadings of ~ 50 μmol g^{-1} (Table 2.4).[93, 132]



DhbtOH **205**

Table 2.4 Support loadings determined following coupling of 3'-*O*-succinates (**200-204**) to LCAA-CPG.

Derivatised support	Support loading μmol g^{-1}
206	50
207	53
208	54
209	49
210	38

Similarly, the free acid function of **200** was coupled to the amino alkyl chain of LCAA-CPG to give **210** (Scheme 2.23). The mechanism presumably involved attachment of the 3'-*O*-succinate (**200**), firstly to DCC, followed by transformation of this adduct into an active ester by reaction with DhbtOH with concomitant displacement of DCU. Reaction of the active ester with the amino function of LCAA-CPG afforded the derivatised support (**210**). The load of $5^{m}C$ nucleoside on the support (**210**) was determined by trityl assay to be 38 μmol g^{-1}, a satisfactory loading, suitable for 1 μmol scale DNA synthesis. Unreacted amino sites on the derivatised support (**210**) were capped by reaction with Ac_2O in the presence of DMAP during 1 h.[76] This solid support derivatisation method has subsequently been used successfully by the Aston group to form the hydroxyhexyl derivatised support (**99**) (Scheme 2.1) with an excellent loading of 68 μmol g^{-1} which was subsequently used in an improved synthesis of SLCPG (**95**) (Section 2.2.1).[93]

2.12 Conclusion

A novel method has been developed using SLCPG support (**95**) which allows for the rapid synthesis of fully deprotected, support-bound sequences containing an amino alkyl tether at the 5'-end. Following derivatisation of the amino function with a suitable ligand, the oligonucleotide-conjugate may be cleaved from the support under neutral conditions using fluoride or, for the first time, using mild acidic conditions. The novel use of acidic conditions greatly extends the repertoire of base-labile ligands which may conjugated to DNA.

In this work, modification of the 5-amino function using **119** and **120** resulted in a significant shift of HPLC retention time compared with the underivatised sequence, allowing for easy HPLC purification. As a result, there was no need for multiple HPLC purification as is standard procedure with others. An advantage of the procedure is that the fully deprotected support- bound amino-linker sequence may be stored and used as necessary allowing isolation of purified conjugates in a matter of a few hours. Future optimisation of the procedure, particularly the Bu^tNH_2 step, may allow for more rapid synthesis, deprotection, conjugate formation, product cleavage and isolation of conjugates providing a significant improvement on existing procedures.

Work towards a similar procedure for the solid-phase synthesis of internal oligonucleotide conjugates of the drugs (**1**) and (**2**) was also carried out. Sequence (**158**) was synthesised using phosphoramidite (**157**) to provide an oligonucleotide containing an AOC-protected aminopentyl linker at the N4 of a single cytosine base. DNA synthesis using phosphoramidite (**157**) and SLCPG (**95**), may provide a means of synthesising internal oligonucleotide conjugates of the drugs (**1**) and (**2**).

The triplex-forming ability of the TFO-conjugates (**132**) and (**134**) was assessed. They were found to form triplexes which were equally as stable as that formed using the unmodified TFO (**27**). Attempts were thus made to effect alkylation of the purine target sequence (**30**) using both conjugates (**132**) and (**134**) and the parent drugs 1 and 2. The use of ESMS analysis failed to provide a mass spectrum when the crude, desalted product mixtures analysed. This may have been due to the low concentration of the analytes or due to traces of salt in the samples which may have adversely affected the low level analysis. Future work in this area could involve the use of an alternative method such as the piperidine method and gel electrophoresis (Section 1.8) to determine the site of alkylation of the 5'-end ^{32}P-radiolabelled purine target following incubation with **132**, **134**, **1**, or **2**. The gel retardation assay (Section 1.8) could be used to isolate any cross-linked products arising from the use of conjugate (**132**) or the parent drug (**1**).

The BPA^t-protected nucleoside (**199**)was synthesised and attached to LCAA-CPG using the 3'-*O*-succinate (**200**). Future work in this area could involve derivatistation of SLCPG support with the 3'-*O*-succinate (**200**) and elaboration of nucleoside **199** to its phosphoramidite form for use in the DNA synthesis of TFO-conjugates of **1** and **2** whose triplex-forming and alkylating ability could be tested at physiological pH.

Chapter 3

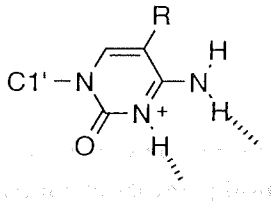
Triplex-Forming Antigene Oligonucleotides

3.1 Introduction and aims

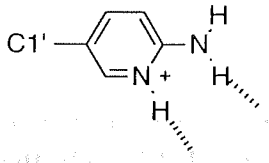
Much effort has been devoted in the last decade to overcoming the pH dependance (Section 1.5.4) of triplexation through the use of various pyrimidine and purine replacements for C in TFOs. It was the aim of this project to enhance the stability of triple helices at physiological pH through the use of purine TFOs containing modified purine bases. The modified bases chosen for analysis were isoguanine (isoG) and two, substituted derivatives of A. The study of isoG in the context of triplexation is the first of its kind and the literature precedents in this area focusing on other bases and analogues are discussed below.

3.2 Pyrimidines and analogues as replacements for C^+

3.2.1 5-Methyl-2'-deoxycytosine



C^+ R = H
5^mC R = CH_3

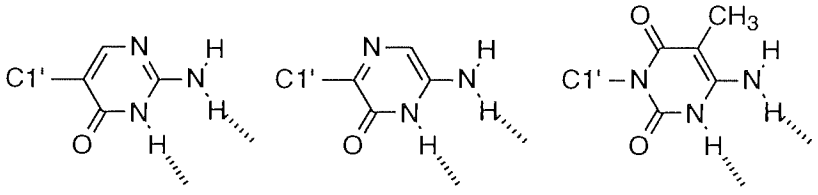


AmP

The use of 5^mC[40, 133] in place of C in the triplex forming strand was found to enhance the stability of triple helices at neutral pH. This stabilisation has been attributed, in part, to the increased basicity of 5^mC (pK_a of 5-methyl-2'-deoxycytidine = 4.5 units)[34] compared with dC (pK_a of 2'-deoxycytidine = 4.3 units)[134] due to the positive inductive effect of the 5-methyl substituent, which leads to more extensive N3 protonation of the base at neutral pH.[39] Also, the positively charged 5^mC residues in the TFO assist in limiting destabilisation due to negative charge repulsion between the phosphate backbones of the TFO and the duplex target.[40] Electrostatic attraction of the positively charged base to the negatively charged phosphate backbone leads to an apparent rise in pK_a of the base. However, thermodynamic analysis[29] has indicated that the enhanced stability was mainly entropic in origin. This was explained by considering the effect of the position of the 5-methyl group of 5^mC in the duplex major groove. It is thought that this hydrophobic group causes release of water molecules of hydration into the bulk

solution, a factor contributing to favourable entropy change. More recently, an even more basic deoxycytidine analogue AmP[134, 135] (pK_a = 6.26)[134] has been used to target G of a G.C base pair. It was found that triplexes containing AmP as a C replacement in the TFO, were more stable than those containing 5^mC at pH 7.5 due to the enhanced basicity of AmP relative to cytidine (pK_a = 4.3).[134] It is thought that triplexes consisting of contiguous runs of 5^mC^+*G.C triplets are energetically unfavourable due to positive charge repulsions between adjacent protonated 5^mC bases.[136] This has led to much interest in the use of synthetic nucleosides that display the hydrogen-bonding pattern of protonated cytosine.

3.2.2 Synthetic pyrimidine analogues as C^+ mimics



The bases isoC[137, 138], pyDDA[139], and m5oxC[41] each possess suitably arranged hydrogen-bond-donating groups which are capable of forming Hoogsteen base pairs with G of a G.C base pair at physiological pH. Importantly, each X*G.C (X = isoC, pyDDA or m5oxC) triplet is isomorphous with the T*A.T triplet (Figure 1.5, Section 1.5.4). Xiang et al[41] found that TFOs containing non-contiguous m5oxC residues formed triplexes which were equally well stabilised at pH values 6.4 and 8.0. However, low triplex stability was observed when five contiguous m5oxC residues were present in the TFO. It was suggested that the 6-carbonyl group of m5oxC may have resulted in poor base-stacking of the m5oxC residues, thereby destabilising the triplex. When $(isoC)_6$ was used as a TFO to target $(G.C)_6$ at pH 7.5, the triplex T_m was ~ 22 °C higher than that when $(5^mC)_6$ was used as the TFO, clearly illustrating the destabilising effect of positive charge repulsion in the $(5^mC)_6$ sequence.[138] The isoC base was more suitable than 5^mC^+ for targeting contiguous G.C doublets. The use of a single pyDDA[139] residue (as its 2'-OMe riboside) in a 15 base TFO has also been shown to support triple helix formation over the pH range 6.3 - 8.0. At five pH values across this range, the T_m of the triplex varied by only ~ 2 °C, illustrating the pH independent manner of triplex formation. When C was used in place of pyDDA, the T_m of the triplex decreased by ~ 10 °C as the pH increased.

3.3 Synthetic purine analogues as replacements for C^+

A number of purine replacements for C^+ has been reported in the recent literature. Each of the bases G^{N7},[140, 141] P1[136, 142] and the two adenosine derivatives 8-oxo-A[33] and 8-oxo-mA[143] have been found to specifically bind G.C base pairs (Figure 3.1). In the latter two cases, the bases adopt a *syn* conformation which allows for triplex formation. Krawczyk et al[143] compared the stability of a 15 base TFO containing either 5^mC or 8-oxo-mA as C replacements at pHs 6.4 and 7.6. Triplex formation was evident at pH 6.4 when either 8-oxo-mA or 5^mC were used, however, at pH 7.6, triplexation was only observed for the TFO containing 8-oxo-mA. Also, when a single T in the TFO was replaced with 8-oxo-mA, no evidence of triplexation was found, thus illustrating the specificity of the 8-oxo-mA*G.C interaction. Triplexes of equal stability were formed at pH values of 7.0 and 8.0 (T_m = 22 °C) using a 15 base TFO containing 3 non-contiguous 8-oxo-A residues[139] whereas the triplex formed by the control C^+-containing TFO decreased in stability as the pH increased from 7.0 (T_m = 28 °C) to 8.0 (T_m = 17 °C). The use of P1[136, 142] in a TFO allowed the sequence specific binding of duplex DNA containing six contiguous G.C base pairs at physiological pH (7.4) which was not possible using the corresponding C^+-containing TFO. Interestingly, specificity for G.C base pairs was lost when P2 was used in place of P1. Explanations for this observation were that either the sugar-phosphaste backbone of the third strand was highly distorted compared with P1-containing triplexes and/or that the methyl group of P2 disfavoured the anti conformation which disrupted the desired P2*G.C hydrogen bonding pattern.

The binding affinity of G^{N7} for G.C base pairs was found to be similar to that of 5^mC when used as a replacement for a single 5^mC base in a 15 base 5^mC and T containing TFO at pH 7.0.[140] G^{N7} was also found to have the greatest specificity for the G.C base pair. Later work by Brunar et al[141] revealed that, at pH values of 7.0 and 7.5, TFOs containing 5^mC bound isolated G.C doublets with higher affinity than those containing G^{N7}, whereas the affinities were reversed where the target duplex contained contiguous G.C base pairs. The first observation was thought to be due to the non-isomorphous nature of the G^{N7}*G.C and T*A.T triplets. Thus many A/G or G/A junctions, ten in this case, in the target duplex would lead to sugar-phosphate backbone distortions in the third strand of the triplex and so to destabilisation. The second observation was attributed to charge repulsion between adjacent 5^mC^+ bases. This work highlighted the effect of non-isomorphous triads in the triplex-forming strand. This problem has been partially alleviated by using TFOs consisting entirely of purine bases.

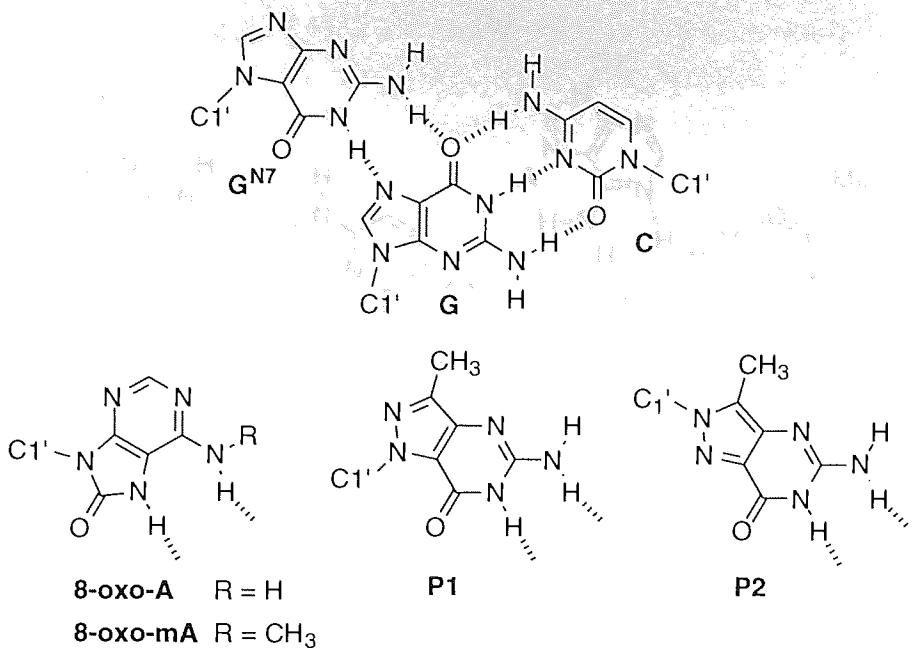
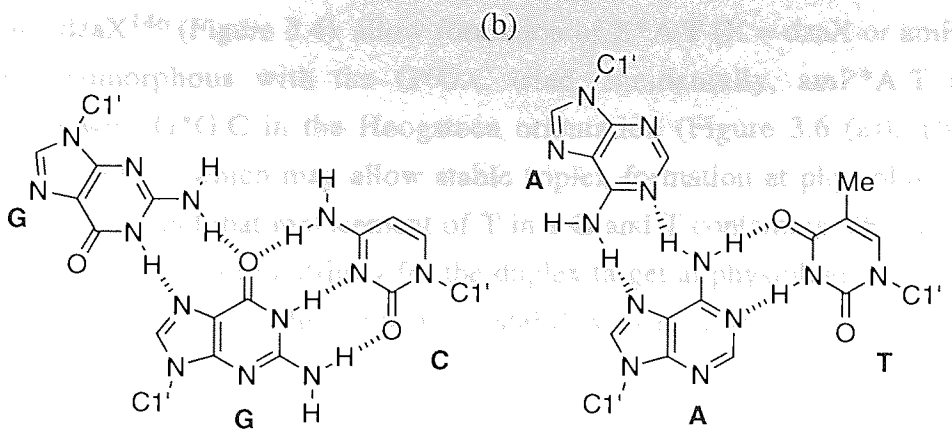


Figure 3.1 Binding scheme for the synthetic purine analogues G^{N7}, 8-oxo-A, 8-oxo-mA, P1 and P2 with the G.C base pair (G^{N7}*G.C base triplet shown).

3.4 Purine-containing TFOs

The natural bases, G and A are known[144] to specifically bind G.C and A.T doublets respectively, forming the triads G*G.C and A*A.T shown in Figure 3.2. Faucon et al[35] used two different 17 base TFOs to target a purine stretch of duplex DNA at pH 7.2. The TFOs used were G and A or G and T containing sequences. Although G and T containing TFOs had been previously used to form triple helices, Faucon et al did not observe triplextation using this oligonucleotide. This was attributed to the eleven G-T or T-G steps within the 17 base oligonucleotide. Moving from a bicyclic purine to a monocyclic pyrimidine would have caused much distortion of the third strand. Hélène et al pointed out that G, T TFOs which had been used elsewhere[145] to form stable triplexes were usually G-rich rather than T-rich and, more importantly, that the G bases were clustered thereby reducing the number of unfavourable steps in the TFO. The G and A containing TFO displayed high affinity and specificity for the duplex target. This is interesting considering that the G*G.C and A*A.T triads are non isomorphous. Clearly, they are closer to isomorphism than the G*G.C and T*A.T triads where monocyclic and bicyclic bases exist in the third strand. Other workers[146] have also noted the lack of binding of G and T containing TFOs and suggested that charge repulsion between the phosphodiesters of the third strand and those of the duplex strands was a factor.



Figures 3.2 Binding schemes for the base triplets: (a) G*G.C and (b) A*AT.

3.4.1 Guanine tetrad formation

Oligonucleotides containing G-rich tracts can adopt a planar tetrameric[147] G-quartet structure which may be composed of inter- or intra-molecular four stranded helices stabilised by mutual bidentate H-bonding. Under physiological salt concentrations, the structure is stabilised by co-ordination of the O6 atoms of the four G bases to a K^+ ion bound between two planes of G quartets (Figure 3.3). Such structures reduce the quantity of TFO available for triplex formation as well as causing interference with thermal denaturation analysis.

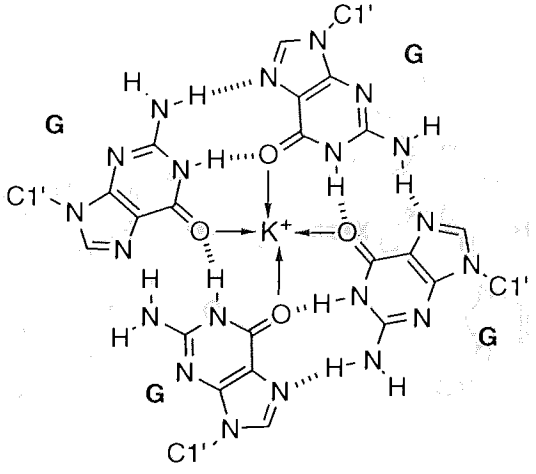
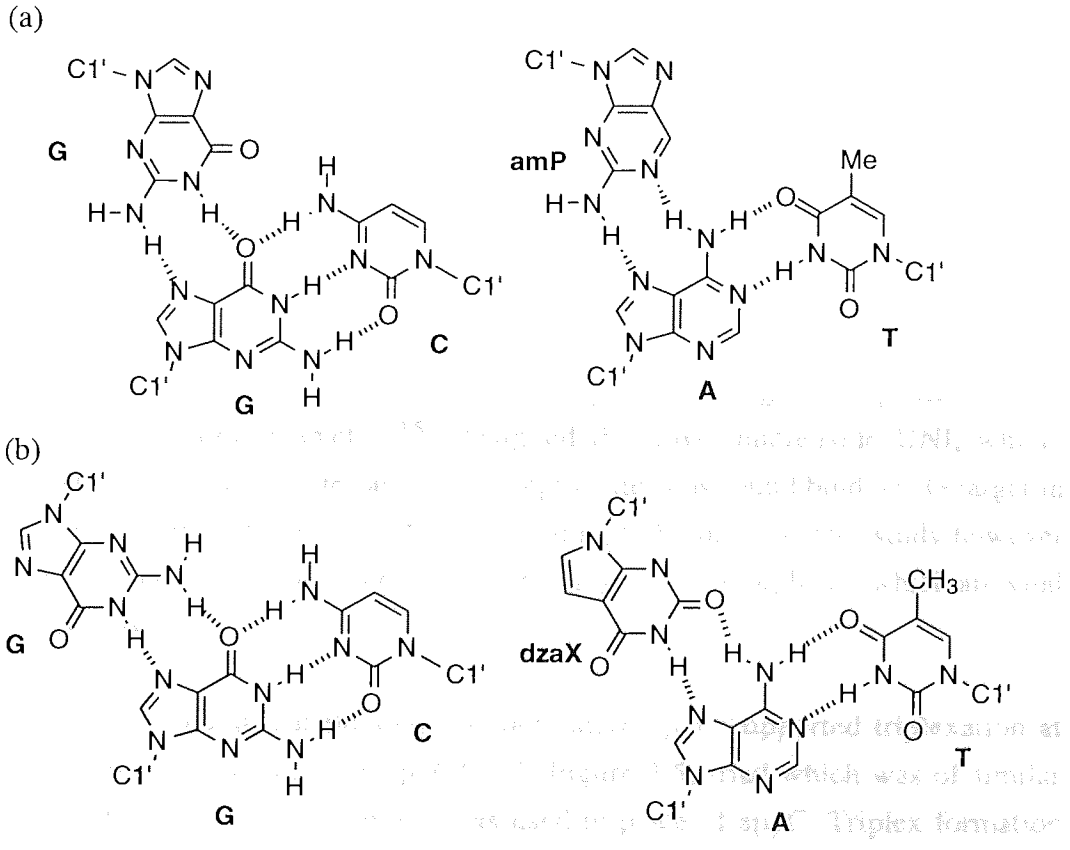


Figure 3.3 Possible structure of a G-tetramer.

3.4.2 Towards isomorphous purine TFOs

Attempts have been made to create TFOs containing isomorphous purine*purine.pyrimidine base triplets. When used as replacements for A, the bases

amP[148] and dzaX[146] (Figure 3.4) allow formation of X*A.T (X = dzaX or amP) triads which are isomorphous with the G*G.C triad. Incidentally, amP*A.T is only isomorphous with G*G.C in the Hoogsteen orientation (Figure 3.6 (a)). These are promising candidates which may allow stable triplex formation at physiological pH. Milligan et al[146] found that replacement of T in a G and T containing TFO with dzaX caused a 100-fold increase in affinity for the duplex target at physiological pH and salt concentrations at 37 °C. This enhanced stability was attributed in part to the isomorphous nature of the dzaX*A.T and G*G.C triads. It was also thought that a decrease in repulsive forces between the phosphate backbone of the third strand (dzaX, G) and the phosphate backbone of the duplex target would have led to enhanced stability.

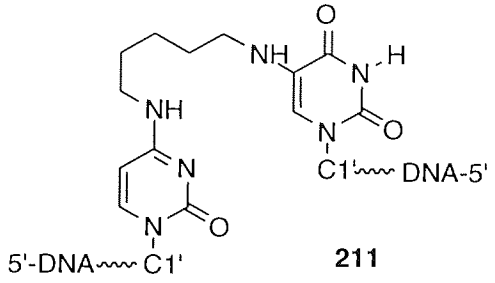


Figures 3.4 Binding schemes for the isomorphous sets of base triplets: (a) G*G.C (Hoogsteen orientation) and amP*A.T and (b) G*G.C (reversed Hoogsteen orientation) and dzaX*A.T.

3.5 Alternate strand triplex formation

Attempts to extend the repertoire of TFO sequences have involved targeting of oligopurine tracts on alternating strands using a single TFO.[149, 150] If the DNA target is

made up of short oligopurine and oligopyrmidine tracts, then a continuous purine tract runs along the duplex, but switches strand at the oligopurine-oligopyrimidine junctions. A series of short TFOs, linked appropriately, would allow the third strand to zig-zag across the duplex major groove to allow binding of the switching purine target. This has been achieved by Zhou et al[151] where two third strand TFOs were joined together by a pentamethylene linker (**211**) which was used to target two oligopurine segments of duplex DNA which alternated on the two DNA strands.



211

3.6 Tridentate bases

Generally, homopurine tracts of DNA are targeted with purine and pyrimidine TFOs since more than one H-bond may be made between the bases of the TFO and the purine target. However, interruption of purine targets by pyrimidine bases poses a problem since, although known, G*T.A and T*C.G triads form by means of only one H-bond and are thus of intermediate stability.[152] In order to expand the range of sequences which may be targeted, Zimmerman et al[153] designed the novel nucleoside UNI, which, containing two H-bond donor sites and one acceptor site, was found bind a C.G target in organic solution ($CHCl_3$) to give UNI*C.G according to Figure 3.5. This study however did not take into account factors such as base stacking and isomorphism, which are vital for triplex stability.

Huang et al[154, 155] found that the cytosine derivative apyC supported triplexation at neutral pH through formation of an apyC*C.G (Figure 3.5) triad which was of similar stability to the triplex formed where C^+ was used in place of apyC. Triplex formation was eliminated when **212** (Figure 3.5) was used in place of apyC in the TFO, illustrating the importance of the 6-amino group of apyC for recognition of the C.G base pair.

It was suggested that, in addition to H-bonding, the pyridine ring of apyC could participate in stabilising, stacking interactions or alternatively intercalate into the duplex. The A.T base pair and to a lesser extent the T.A and G.C base pairs were also recognised by apyC.

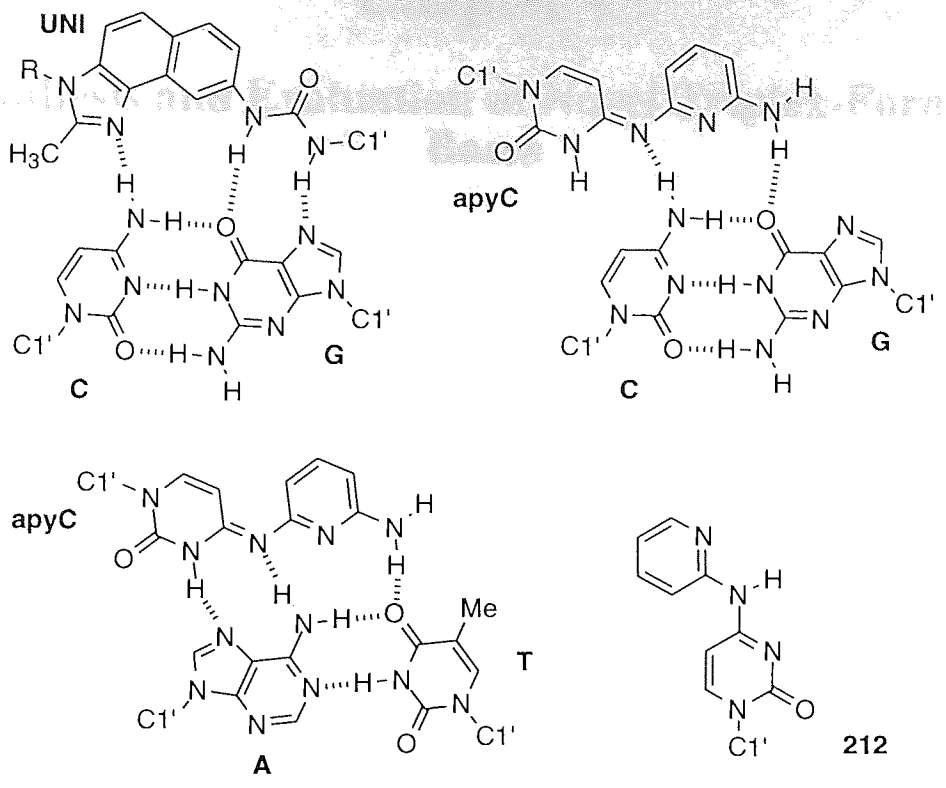


Figure 3.5 Binding scheme for the base triplets UNI*C.G, apyC*C.G and apyC*A.T.

Interestingly, the triplex containing apyC*A.T (Figure 3.5) was found to be more stable than that containing T*A.T. The enhanced stability was attributed to the additional H-bond of the apyC*A.T triad compared with the T*A.T triad (Figure 1.5, Section 1.5.4).

Chapter 4

Synthesis and Evaluation of Novel Triplex-Forming Bases

4.1 Introduction

4.1.1 Aims

One aim of this work was to synthesise 2'-deoxyisoguanosine (**217**) (isoG) via the 2-chloro-2'-deoxyadenosine intermediate (**214**). Also of importance was the incorporation of the isoG base into pentose DNA to study:

1. Whether the N3 tautomer of isoG could be used as an replacement for A in a TFO to target an A.T base pair.

2. Its effect on triple helix formation/stability when used as a replacement for G in a G- and A-containing TFO.

4.1.2 The suitability of isoG as a triplex-forming base

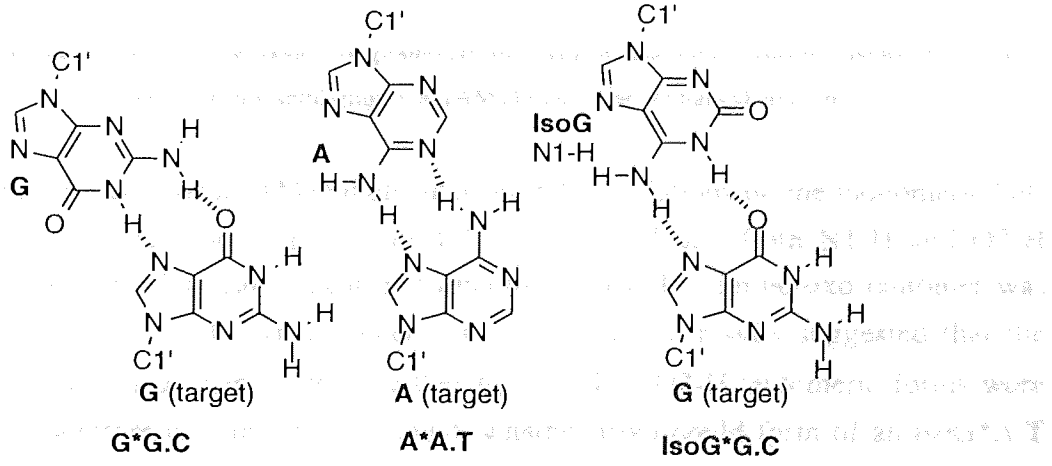


Figure 4.1 Binding schemes for the base triplets: G*G.C, A*A.T and isoG*A.T. The pyrimidine bases of each base triplet have been omitted for clarity.

In G and A containing TFOs, A binds A.T and G binds G.C by means of Reversed-Hoogsteen hydrogen bonds (Figure 3.2). Importantly, the A*A.T and G*G.C triplets are not isomorphous as shown in Figure 4.1 (only the purine bases of each triplet are shown for clarity). Isomorphism of base triplets results in a uniform deoxyribose-phosphate

backbone, which provides triplex stabilisation.[34] It has previously been observed that a reduction in the number of unfavourable 'steps' in a TFO leads to greater triplex stability.[35] Therefore, since the isoG*G.C triplet is isomorphous with the A*A.T triplet (Figure 4.1), it was considered that replacement of a G base with isoG, flanked on either side by A bases, in a TFO would lead to stabilisation of the triplex due to removal of some backbone distortion.

4.1.3 Tautomeric forms of the isoG base

IsoG is known to exist readily in either of the N1-H or O2-H tautomeric forms shown in Figure 4.2, depending on the solvent dielectric constant. In aqueous media, the N1 tautomer is prevalent, contributing ~90 % of the total, while the O2-H tautomer constitutes the remaining 10%.[156]

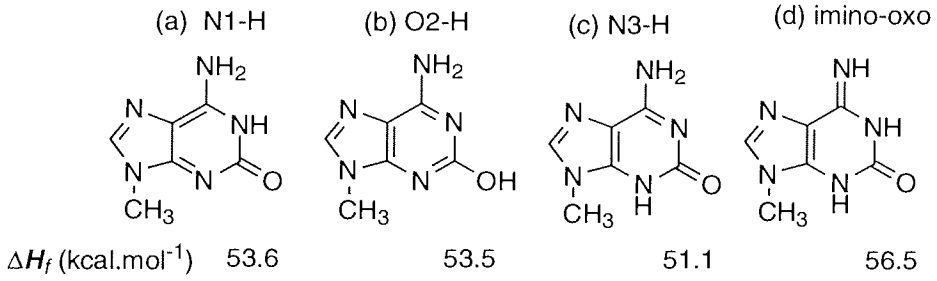


Figure 4.2 Tautomeric forms of the model compound N9-methyl isoG and their respective heats of formation (ΔH_f) derived using semi-empirical (AM1) molecular orbital calculations.

In this work, semi-empirical (AM1) molecular orbital calculations on the monomeric N9-methyl isoG tautomers shown in Figure 4.2 were carried out. Both N1-H and O2-H tautomers were near identical in heat of formation (ΔH_f). The imino-oxo tautomer was highest in ΔH_f while the N3-H tautomer was lowest. These results suggested that the prospects of targeting A using isoG in either the N3-H or O2-H tautomeric forms were good. It was therefore of interest to establish whether isoG could form of an isoG*A.T triplet in either tautomeric form (N3-H or O2-H) (Figure 4.3) and to compare its stability with that of the A*A.T triplet. Both of the isoG*A.T triplets shown in Figure 4.3 are isomorphous with the A*A.T triplet.

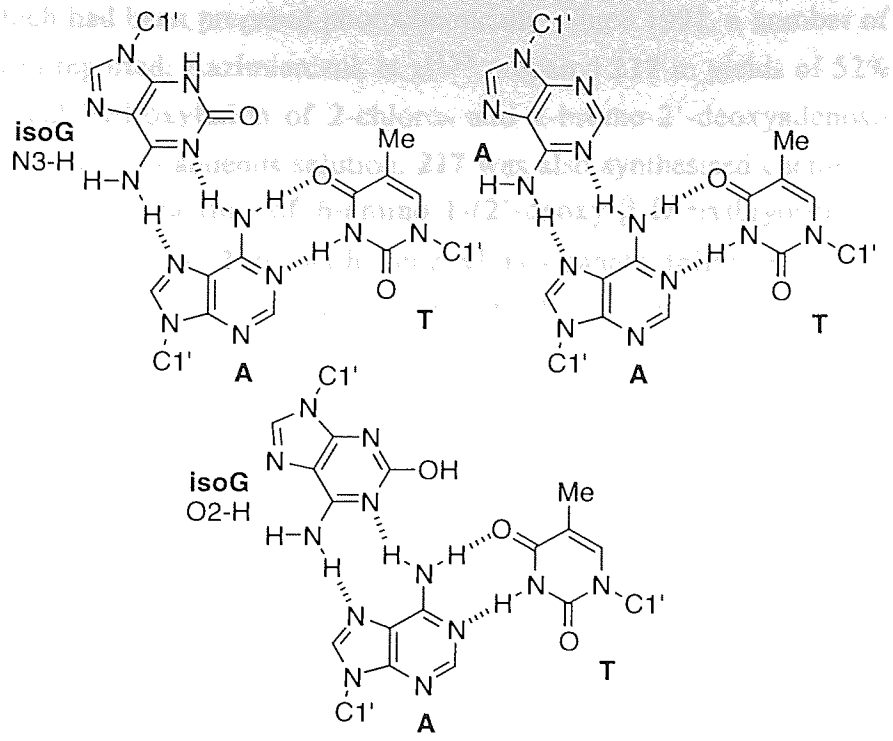
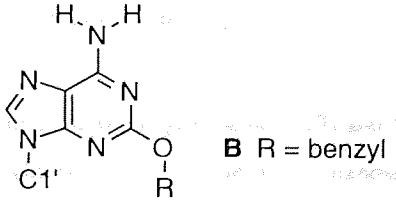


Figure 4.3 Binding schemes for the base triplets: isoG (N3-H)*A.T, A*AT. and isoG (O2-H)*A.T.

4.1.4 Analogues of the adenine base

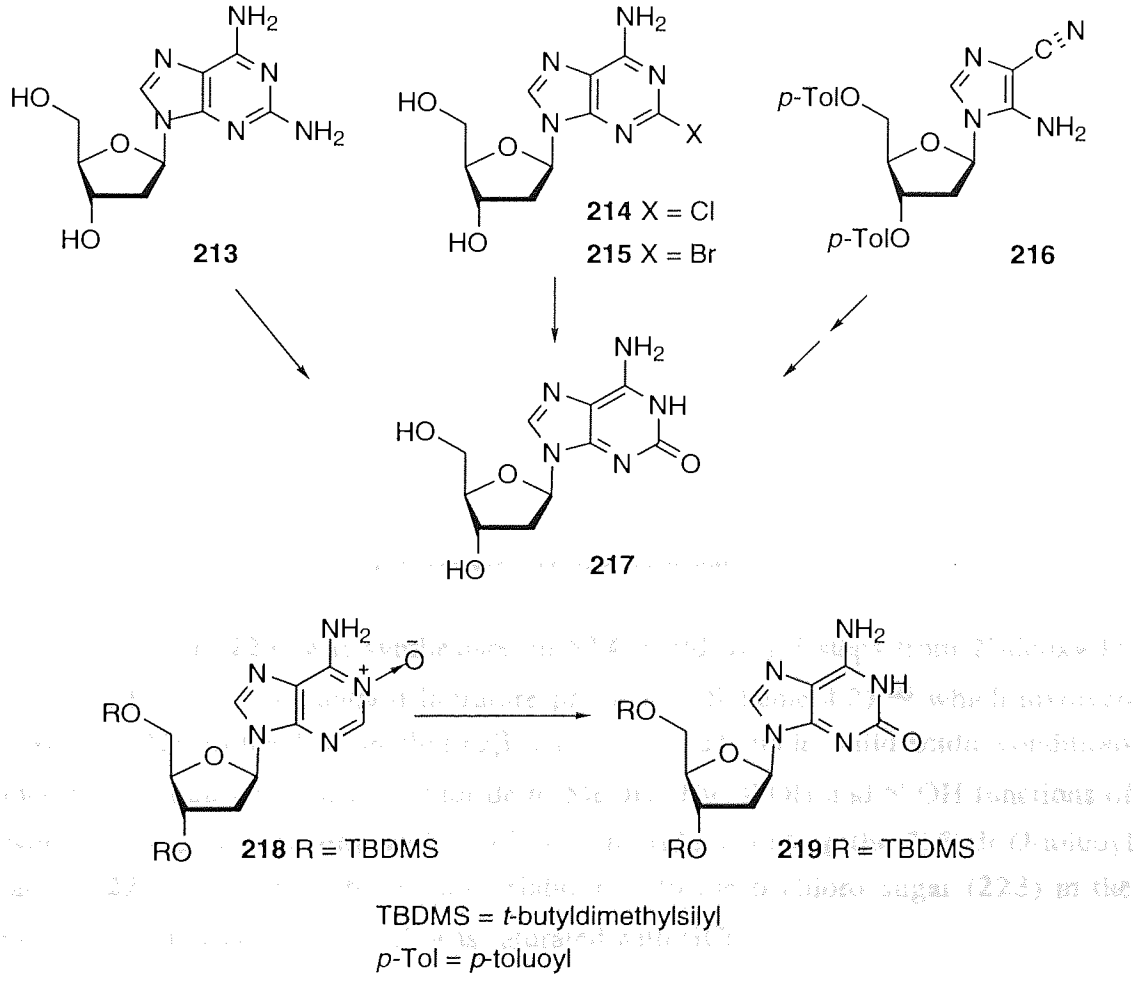
Another area of interest in this work was prompted by the findings of Groebke et al.[157] These workers noted significant increases in the stability of homo-DNA duplexes containing a 2-*O*-benzyl derivative (**B**) of adenine. Homo-DNA are linear where the pentose ring of natural DNA is replaced by the six-membered ring of 2',3'-dideoxyglucose. Therefore, it was of interest to observe the effect of the same A analogue on triple helix formation and stability when incorporated into natural pentose DNA to target an A.T base pair. It was anticipated that the benzyl substituent would be accommodated in the duplex major groove, possibly resulting in stabilisation of the triple helix.



4.2 Literature precedents of the synthesis of 2'-deoxyisoguanosine (217)

Until 1991, there were no reported syntheses of 2'-deoxyisoguanosine (**217**) in the literature. Up to that time, only the 5'-mono- and 5'-triphosphate derivatives of (**217**) were

known, which had been prepared photochemically. Since 1991, a number of syntheses of **217** has been reported. Kazimierczuk et al[158] prepared **217** in yields of 52% and 53% by photochemical hydroxylation of 2-chloro- and 2-bromo-2'-deoxyadenosine (**214**) and (**215**), respectively, in aqueous solution. **217** was also synthesised chemically, in a lower yield of 43%, by reaction of 6-amino-1-(2'-deoxy-β-D-erythryopentofuranosyl)-1*H*-imidazoyl-4-carbonitrile (**216**) with benzoyl isocyanate followed by treatment with ammonia. A drawback of this method is that it relies on the use of the 2'-deoxyriboside (**216**) which is not commercially available.



Scheme 4.1 Literature routes towards 2'-deoxyisoguanosine (**217**) and the silyl-protected derivative 3',5'-*O*-bis(*t*-butyldimethylsilyl)-2'-deoxyisoguanosine (**219**).

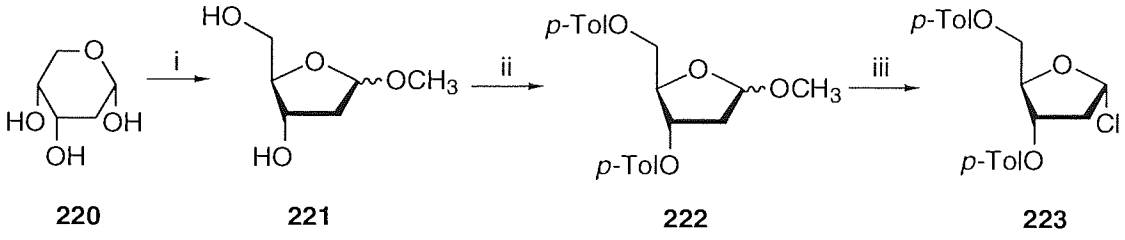
Switzer et al[156] used photolysis to convert 3',5'-*O-bis*(*t*-butyldimethylsilyl)-2'-deoxyadenosine *N*-oxide (**218**) to 3',5'-*O-bis*(*t*-butyldimethylsilyl)-2'-deoxyisoguanosine (**219**) in 60% yield. However, formation of the *N*-oxide intermediate (**218**) from the 3',5'-*O*-bis(*t*-butyldimethylsilyl)-2'-deoxyadenosine starting material using aqueous ethanolic

monopermaleic acid took 2 weeks with a low yield of 39%. Recently,[159] the yield of the **218** has been improved to 80% through the use of monochloroperbenzoic acid in CH_2Cl_2 for the conversion. More recently, Seela et al[160] have synthesised **217** in 42% overall yield by deamination of 2-amino-2'-deoxyadenosine (**213**) (59% yield) which was formed by amination of the silylated 2'-deoxyguanosine starting material (71% yield).

4.3 Synthesis

4.3.1 1-Chloro-2-deoxy-3,5-di-*O*-*p*-toluoyl-α-D-*erythro*-pentofuranose (223)

In this work, 2-chloro-2'-deoxyadenosine[161, 162] (**214**) (Scheme 4.3) was chosen as an intermediate towards the synthesis of both the free nucleoside (**217**) and a phosphoramidite form of **217**. The sugar component of **217** was synthesised as follows.



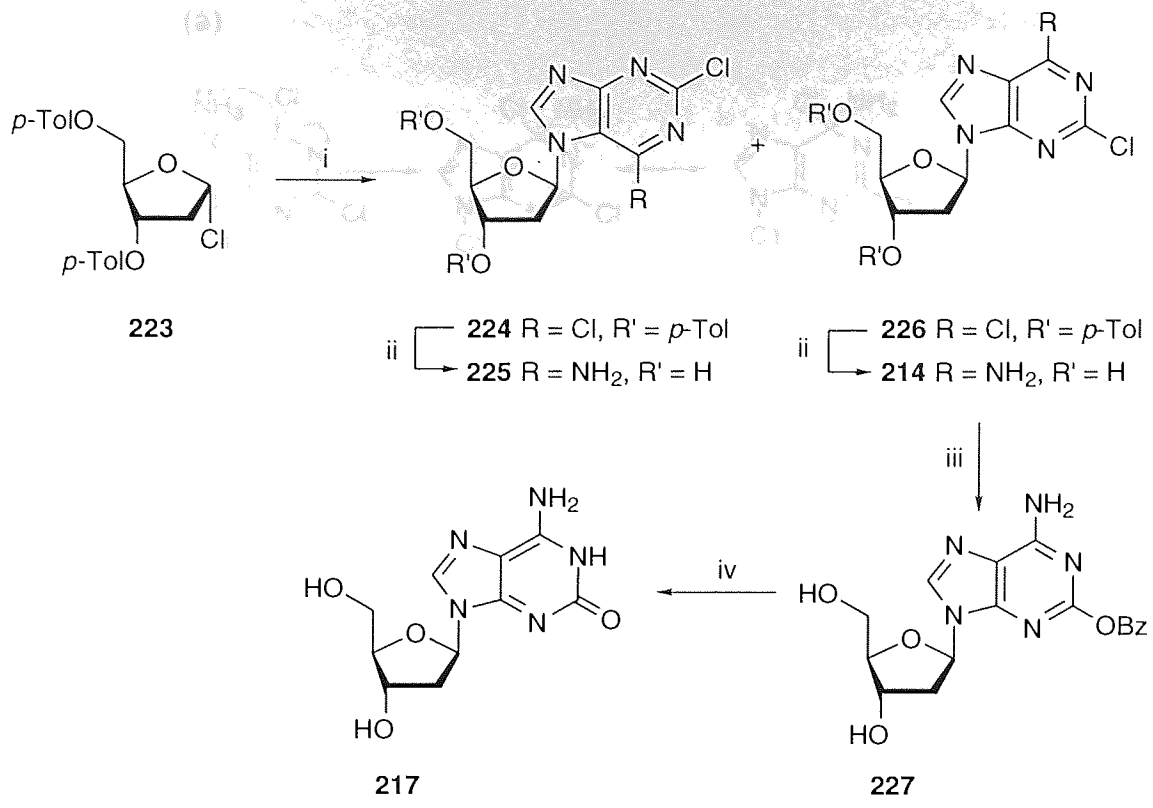
220 **221** **222** **223**

Scheme 4.2 Synthesis of 1-chloro-2'-deoxy-3,5-di-*O*-*p*-toluoyl-α-D-*erythro*-pentofuranose (**223**). i, Acetyl chloride, MeOH, Ag_2CO_3, r.t., 10 min; ii, pyridine, *p*-toluoyl chloride, 0 °C, 24 h 30 min; iii, glacial AcOH, HCl, 0 °C, 30 min.

The α-chloro sugar (**223**) was synthesised in 52% yield over 3 steps from 2'-deoxy-D-ribose (**220**), following a standard literature procedure (Scheme 4.2)[163] which involved conversion of **220** to the 1'-*O*-methyl α/β derivative (**221**) under mild acidic conditions provided by the addition of acetyl chloride to MeOH. The 3'-OH and 5'-OH functions of **221** were protected by reaction with *p*-toluoyl chloride affording the 3',5'-di-*O*-toluoyl derivative **222** which was subsequently elaborated to the α-chloro sugar (**223**) in the presence of glacial acetic acid which was saturated with HCl.

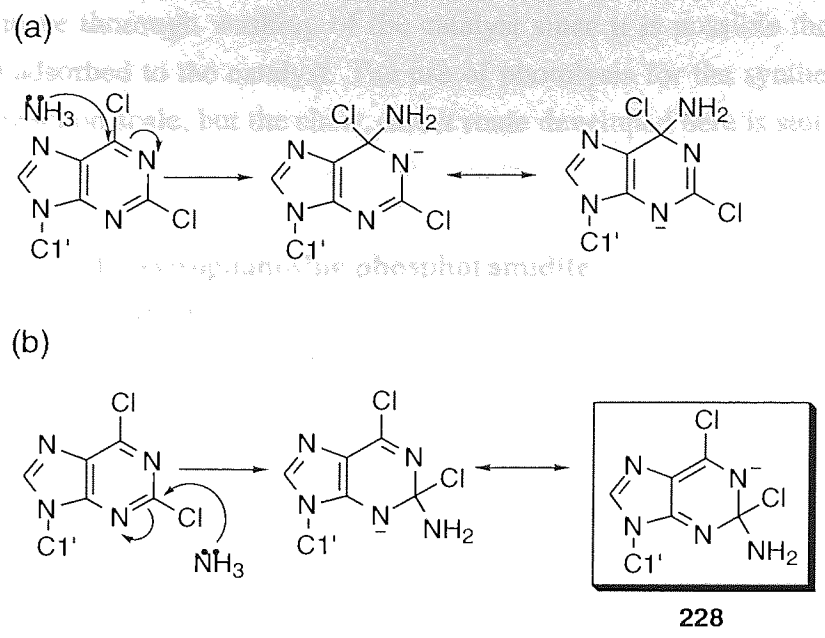
4.3.2 Synthesis of 2-chloro-2'-deoxyadenosine (214)

The sodium salt of 2,6-dichloropurine was generated, *in situ*, by deprotonation of the N9 proton of 2,6-dichloropurine using NaH in MeCN. Addition of **223** to the reaction afforded the desired β forms of the regioisomers **226** (N9) and **224** (N7), as a 2:1 mixture as judged by ^{1}H NMR spectroscopy (Scheme 4.3).



Scheme 4.3 Synthesis of 2'-deoxyisoguanosine (**217**) and its precursors (**227**) and (**214**). i, NaH, 2,6-dichloropurine, MeCN, r.t., 25 h 30 min; ii, MeOH-NH_3 (sat.), 100 °C, 6 h; iii, sodium benzylalkoxide, 85 °C, 12 h; iv, MeOH, 5% Pd-C, H_2, r.t., 6.5 h.

Formation of the N7 isomer resulted from resonance stabilisation of the 2,6-dichloropurine anion. The reason for the preferential formation of **226** over **224** is probably due to steric hindrance caused by the presence of the 6-chloro group resulting in a reduced yield of the N7 isomer **224**. Isomers (**224**) and (**226**) could not be purified by flash chromatography as they were not well separated by TLC using a number of different solvent systems. Therefore, the crude product mixture was used directly in the next stage. The chloro group at the C6 position of **226** is known to be more easily displaced than that at C2.[161] This is explained by the fact that each of the three resonance structures which may be drawn for the intermediate anion arising from attack by ammonia at C6 involve retention of the stable aromatic 5-membered ring-system (Scheme 4.4 (a)). However, in the case of attack by NH_3 at C2, one of the three resonance structures (**228**) (Scheme 4.4 (b)) which may be drawn for the intermediate anion involves disruption of the aromatic 5-membered ring-system.



Scheme 4.4 Resonance structures of the anion generated from attack by NH_3 at either: (a) C6 or (b) C2 of **226**.

Thus the mixture of **224** and **226** was treated with methanolic ammonia solution during 6 hours, while heating at 100 °C, which afforded the isomers **214** and **225** in yields of 14% and 3% over two steps, following flash chromatographic purification (Scheme 4.3). No further use was made of the N7 isomer **225**. Towards **217**, insertion of a protected form of the 2-oxo function was realised by displacement of the C2 chloro group of **214** using benzylalkoxide.

4.3.3 2-*O*-Benzyl-2'-deoxyadenosine (227)

A 50% dispersion of sodium in paraffin oil was heated at 80 °C in benzyl alcohol during 45 min, generating benzyl alkoxide. Addition of the N9 nucleoside (**214**) to the mixture, with continued heating at 85 °C during 12 hours, afforded the 6-amino-2-benzyloxy derivative (**227**) (Scheme 4.3), which was isolated in 71% yield following chromatographic purification.

4.3.4 Synthesis of 2'-deoxyisoguanosine (217)

The benzyl group was removed from **227** by hydrogenolysis using Pd on C (10%) during 6.5 h, yielding 2'-deoxyisoguanosine (**217**) in 53% yield (Scheme 4.3). **217** was thus synthesised by a novel route from the 2-chloro-6-amino nucleoside **214**, in 38% overall yield, over 2 steps. This yield was lower than that achieved by Kazimierczuk et al[158] who used photolysis to convert **214** to **217** in one step (52% yield). The yield of **217** could be

improved by more thorough washing of the catalyst since it is possible that some of the product (**217**) adsorbed to the catalyst. The use of photolysis for the synthesis of **217**[156, 158] limits the reaction scale, but the short, direct route developed here is suitable for scale-up.

4.4 Synthesis of 2'-deoxyisoguanosine phosphoramidite

4.4.1 Literature precedents

In order to incorporate isoG into a oligonucleotide sequence, synthesis of a phosphoramidite form of **217** was necessary. A number of phosphoramidites of **217** are known in both the ribose and 2'-deoxyribose series. One particularly important synthesis of an isoG phosphoramidite in the ribose series was carried out by Tuschl et al[164] in 1993.

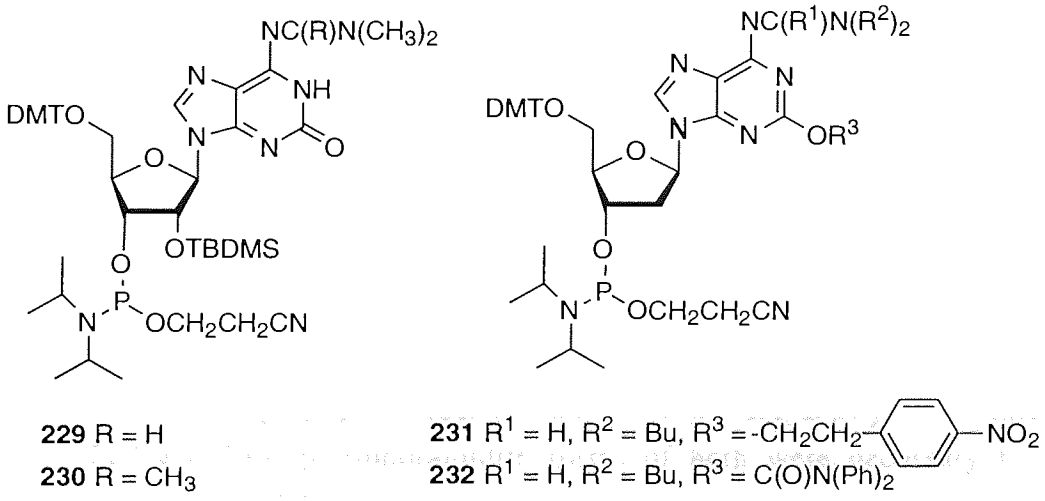


Figure 4.4 Various phosphoramidites monomers which have been used to incorporate isoG into RNA (**229**, **230**) and DNA (**231**, **232**) sequences.

These workers incorporated the isoG base into synthetic RNA using **229** (Figure 4.4). However, it was found that once the base had been incorporated at the 3'-end of a sequence, further elongation of the oligomer was unsuccessful. This was later attributed to the loss of the formamidine amino-protecting group during oligonucleotide synthesis.[165] The exposed amino function was then available for reaction with an incoming phosphoramidite, resulting in unwanted, 'branched' oligonucleotides.[165] Ng et al[165] later synthesised a variety of amino protected isoguanine nucleosides including phenoxy acetyl, benzoyl and isobutyryl derivatives, all of which were isolated in very poor yields of 10%, 27% and 34% respectively. However, the acetamidine protected phosphoramidite **230** (Figure 4.4) was used successfully to incorporate the isoG base at three internal positions within a ribozyme. Similar to the 6-oxo function of guanine,[166] the 2-oxo function of **217**

is reactive enough to be phosphitylated by an incoming phosphoramidite during DNA synthesis,[167] leading to the formation of 'branched oligonucleotide' impurities and poor yields. As a result, a number of 2-oxo protecting groups have been employed to prevent these unwanted side-reactions. In the 2'-deoxyribose series, improved DNA synthesis has been realised through the use of 4-nitrophenethyl[156], and more recently, diphenyl carbamoyl[159, 168] protected phosphoramidites (**231**) and (**232**) respectively (Figure 4.4).

In order to incorporate isoG into a oligonucleotide sequence, a suitable 2-oxo protecting group was required. The use of 2-*O*-benzyl derivative **227** for this purpose was considered inappropriate as binding of oligonucleotides to the catalyst is known to occur.[157] The allyl group has been used successfully for protection of the 2-oxo group of isoG which allowed generation of the isoG base in hexose DNA[157] following cleavage of the allyl group using Pd(0). This protecting group and was therefore considered suitable for the purpose of incorporating isoG into pentose DNA.

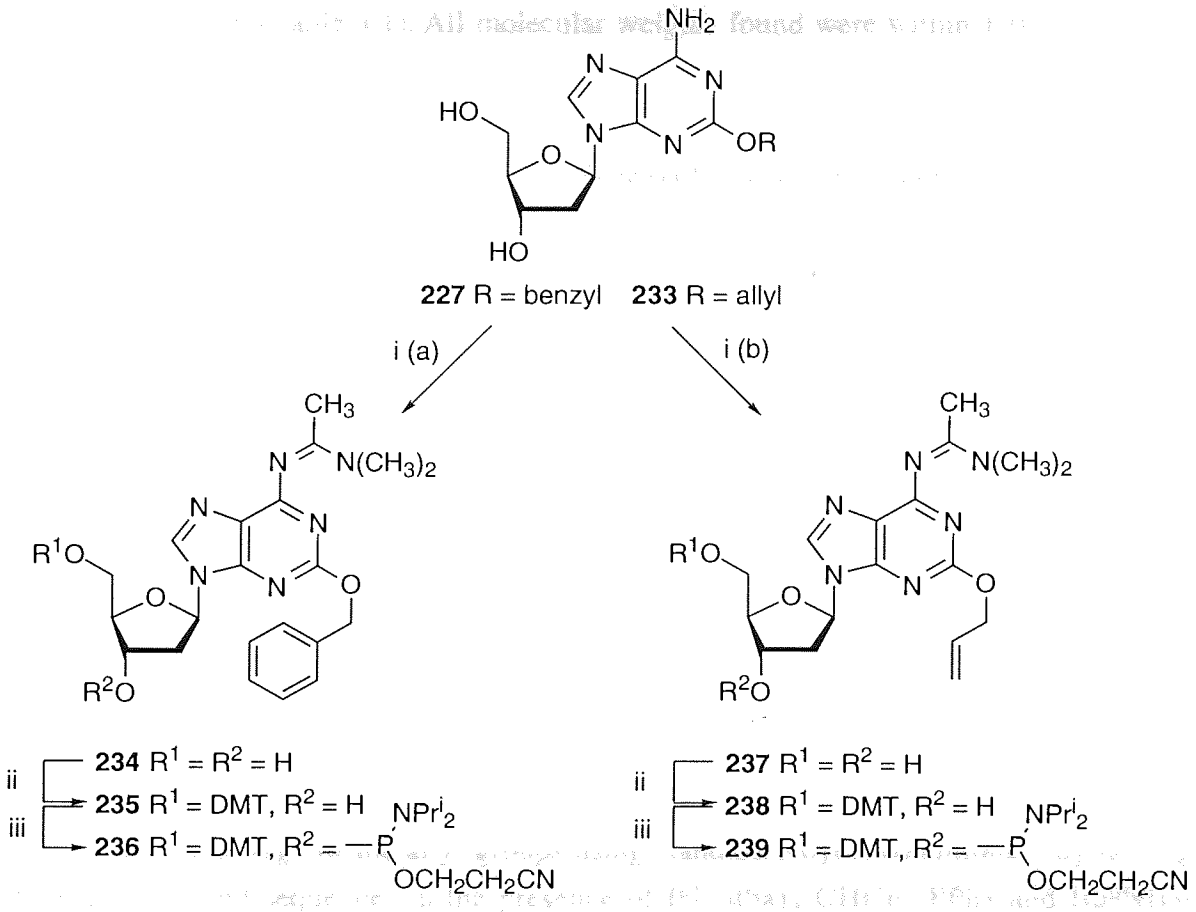
4.4.2 Synthesis of 2-*O*-allyl-2'-deoxyadenosine (233)

The 2-*O*-allyl derivative (**233**) (Scheme 4.5) was synthesised in a similar fashion to the 2-*O*-benzyl derivative (**227**). A solution of sodium allylalkoxide was prepared by reacting a 50% dispersion of sodium in paraffin in allyl alcohol at 0 °C during 1 h. Addition of **214** to the mixture with heating at 88 °C during 12 h, afforded the **238** in 62% yield after purification by chromatography. In order to study the triplexing ability of the adenosine analogues **227** and **233**, phosphoramidite forms of both were necessary to allow incorporation into pentose DNA.

4.4.3 Synthesis of phosphoramidites (236) and (239)

The 2-amino functions of adducts (**227**) and (**233**) were protected by reaction with *N,N* -dimethylacetamide dimethyl acetal in anhydrous methanol during 26 h (40 °C) and 47 h (r.t.) respectively according to a literature procedure[169] (Scheme 4.3). The volatile reagents were removed under low vacuum and the crude product mixtures were used in the next step without further purification. The 5'-hydroxyl functions of **234** and **237** were tritylated using 4,4'-dimethoxytrityl chloride affording **235** and **238** in satisfactory yields of 88% and 68% over 2 steps, respectively, following purification by chromatography. Phosphitylation at the 3'-hydroxyl functions of **235** and **238** using 2-cyanoethyl-*N,N*-diisopropyl chlorophosphoramidite afforded the DNA building block monomers **236** and **239** in yields of 33% and 64% respectively, as mixtures of diastereoisomers. In both cases, precipitation of the chromatographically purified products into hexane was required to

remove an unknown non-phosphoramidite impurity which had an intense signal in the ^{31}P NMR spectrum (~ 14 ppm).



Scheme 4.5 Synthesis of phosphoramidites (**236**) and (**239**). i (a), dry MeOH, *N,N*-dimethylacetamide dimethyl acetal, 40 °C, 26 h; i (b), dry MeOH, *N,N*-dimethylacetamide dimethyl acetal, r.t., 47 h; ii, pyridine, DMTCl, r.t., 12 h for (**235**), 110 min for (**238**); iii, THF, 2-cyanoethyl-*N,N*-diisopropyl chlorophosphoramidite, r.t., 40 min for (**236**) or 1 h for (**239**).

4.5 Solid phase oligonucleotide synthesis

240	3'-d(TTTTCTTTTCCCCCCT)
241	5'-d(AAAAGAAAAGGGGGGA)
242	3'-d(AAAAGAAAAGGGGGGA)

The triplex helix **242*241.240** was chosen as a suitable system with which to evaluate isoG (I) and the adenine derivatives 2-*O*-benzyl (B) and 2-*O*-allyl (L). Synthesis of sequences (**240-247**) was carried out on a 1 μmol scale where phosphoramidites (**236**) and (**239**) were

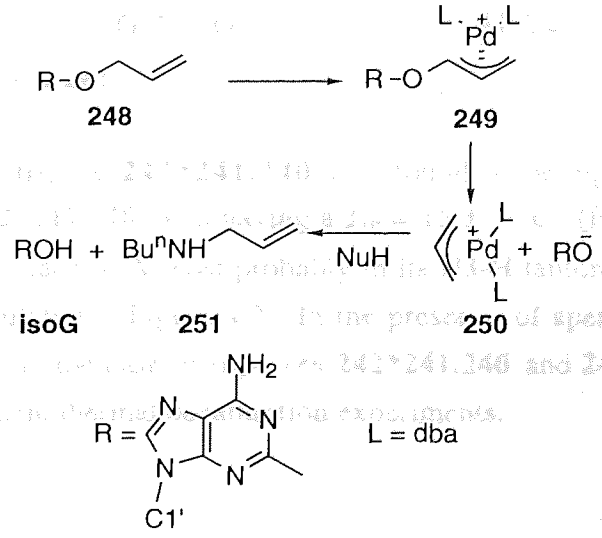
used to insert the 2-*O*-benzyl and 2-*O*-allyl-adenine residues, respectively. Cleavage and deprotection using conc. NH_3 (aq) at 55 °C during 16 h afforded the crude oilgomers which were purified by HPLC and analysed by ESMS in the negative mode in weakly basic ammonia solution (Table 4.1). All molecular weights found were within ± 0.08% of their calculated values.

Table 4.1 HPLC/ESMS analysis of oligonucleotide sequences (**240-247**).

	Sequence	t_R/min	MW Found	MW Calculated
240	3'-d(TTTTCTTTTCCCCCCT)	13.8	4698.8	4697.8
241	5'-d(AAAAGAAAAGGGGGGA)	12.6	5061.4	5058.9
242	3'-d(AAAAGAA**A**AGGGGGGA)	12.4	5060.7	5058.9
243	3'-d(AAAAGAA**I**AGGGGGGA)	12.7	5076.9	5074.9
244	3'-d(AAAAGAA**L**AGGGGGGA)	12.6	5119.7	5115.0
245	3'-d(AAAAGAA**B**AGGGGGGA)	14.2	5168.0	5165.0
246	3'-d(AAAAGAA**G**AGGGGGGA)	12.7	nd	5074.9
247	3'-d(AAAA**I**AAAAGGGGGGA)	12.7	5062.8	5058.9

L = 2-*O*-allyl-A, B = 2-*O*-benzyl-A, I = isoG. .nd = not determined.

The isoG bases were generated in sequences (**243**) and (**247**) prior to treatment with conc. NH_3 (aq), by cleavage of the allyl groups using standard Noyori conditions[86] by heating the support bound sequences in the presence of $Pd_2(dba)_3$-$CHCl_3$, PPh_3 and Bu^nNH_2-HCO_2H in THF at 55 °C during 12 h.



Scheme 4.6 Mechanism of Pd(0)-mediated cleavage of the allyl group from sequences **243** and **247**.

The mechanism of the allyl cleavage reaction is shown in Scheme 4.6. The mechanism involves co-ordination of the electron-rich allyl group of **248** with the complex $Pd_2(dba)_3$-$CHCl_3$ to give the π-allyl palladium intermediate (**250**) via the intermediate **249**. The allyl group of **250** is trapped by reaction with Bu^nNH_2 affording **251** and the isoG base as its O2-H tautomer.

4.6 Triplexation studies of TFOs (242-246) at pH 7.2

Thermal denaturation of the triple helices using the TFOs containing the modified bases was carried out at pH values 7.2 and 6.4 using oligonucleotides at a concentration of 1.5 μmol each in either TRIS buffer (pH 7.2) or PIPES buffer (pH 6.4) .

240 3'-d(TTTTCTTTTCCCCCCT)
241 5'-d(AAAAGAAAAGGGGGGA)
242-246 3'-d(AAAA**Y**AA**X**AGGGGGGA)

Table 4.2 Duplex and triplex melting temperatures (T_m) for triplexes (**242-246**)***241.240**.

Triplex	TFOs (242)-(246)	T_m/°C	T_m/°C	T_m/°C
242*241.240	X = A, Y = G	15 ± 1	-	59.0 ± 0.5
243*241.240	X = IsoG, Y = G	15 ± 1	-	59.0 ± 0.5
242*241.240[a]	X = A, Y = G	23 ± 1	-	59.0 ± 0.5
243*241.240[a]	X = IsoG, Y = G	23 ± 1	-	59.0 ± 0.5
244*241.240	X = L, Y = G	16 ± 2	41 ± 2	59.0 ± 0.5
245*241.240	X = B, Y = G	15 ± 3	45 ± 2	59.0 ± 0.5
246*241.240	X = G, Y = G	-	41 ± 2	59.0 ± 0.5

[a] 0.5 mM spermine added.

The isoG-containing triplex **243*241.240** was found to be equally as stable as the unmodified triplex **242*241.240**, both having a T_m = 15 ± 1 °C (Figure 4.5). Thus, isoG was accommodated in place of A most probably in its N3-H tautomeric form based upon molecular orbital calculations (Figure 4.2). In the presence of spermine, triplex stability was enhanced by 8 °C in the case of triplexes **242*241.240** and **243*241.240**. Spermine was used in all subsequent thermal denaturation experiments.

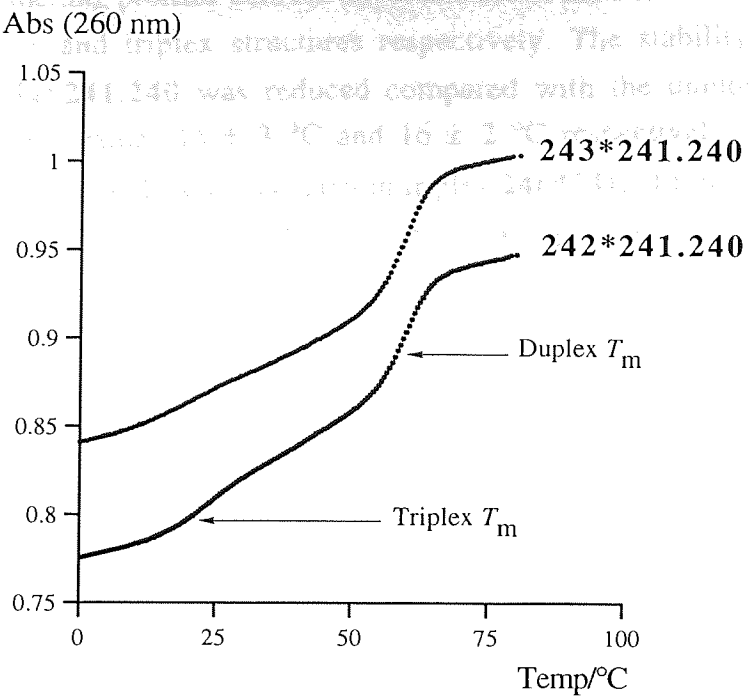


Figure 4.5 Melting profile of triplexes **242*241.240** and **243*241.240**.

However, the duplex T_m = 59.0 ± 0.5 °C was unaffected by the presence of spermine. Other workers[168,170] have observed the formation of tetraplex structures for sequences containing isoG similar to the G-tetraplex structures described in Section 3.4.1. To investigate whether the transitions observed in these experiments were due to tetraplex formation, TFOs (**242**) and (**243**) were each incubated separately in the absence of the duplex target **241.240**. Melting transitions were not observed in either case ruling out the possibility of tetraplex formation.

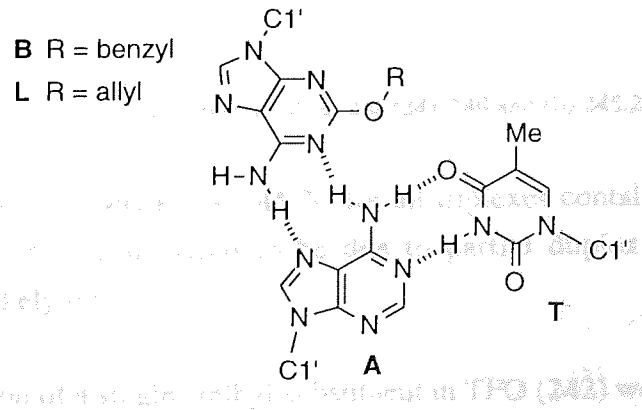


Figure 4.6 Binding schemes for the base triplets: B*A.T, and L*A.T.

When A was replaced at position X by B (Figure 4.7) or L in the TFO, three transitions were observed in the melting profiles with the upper and lower transitions corresponding to melting of the duplex and triplex structures respectively. The stability of triplexes **245*241.240** and **244*241.240** was reduced compared with the unmodified triplex **242*241.240** having T_m values 15 ± 3 °C and 16 ± 2 °C respectively. When A was replaced with G to provide a G*A.T mismatch in triplex **246*241.240**, triplex formation was completely inhibited as evidenced by the lack of an appropriate transition in the melting profile, illustrating the sensitivity of the triplex **242*241.240** to a single mismatch.

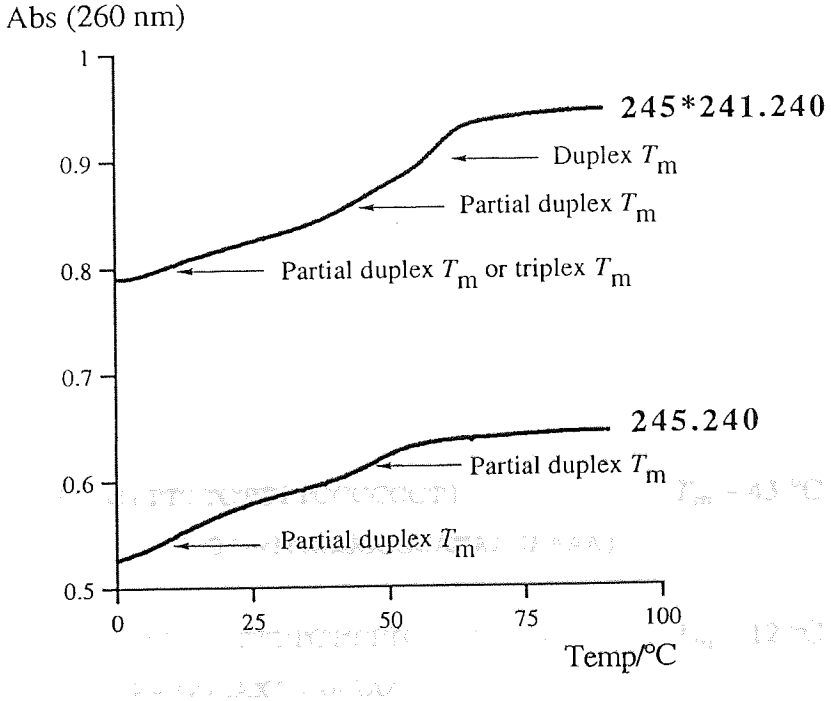


Figure 4.7 Melting profile of (a) **245*241.240** and (b) **245.240**.

A transition was observed between 40-45 °C for all triplexes containing B, L or G. In all cases this transition was considered to be due to partial duplex formation as it was considered very unlikely that:

1. The incorporation of a single aralkyl substituent in TFO (**242**) would enhance the triplex T_m from 23 °C to 40-45 °C.

2. The triplex **246*241.240** containing a G*A.T mismatch would have a T_m which was 18 °C higher than that of the matched triple helix **242*241.240**.

Partial duplex formation could conceivably occur as a result of mismatches between B, L, or G in the TFO strand and the A target of the duplex. When sequences **242-246** were annealed with **240** alone, two transitions were found in each case (Table 4.3).

Table 4.3 Melting temperatures (T_m) observed for duplexes (**242-246**).**240**.

Duplex	Base	T_m/°C	T_m/°C
242.240	**X** = **A**	22 ± 3	43 ± 2
243.240	**X** = IsoG	23 ± 2	43 ± 2
244.240	**X** = L	10 ± 2	44 ± 2
245.240	**X** = B	11 ± 1	46 ± 2
246.240	**X** = G	13 ± 3	40 ± 2

The lower transition T_m ~ 12 °C (X = L, B (Figure 4.7) or G) or ~ 22 °C (X = A or isoG) appeared to be dependent on the structure of X while the upper transition (T_m ~ 43 °C) appeared to independent of the structure of X. Partial duplex structures which correlate with these observations are shown in Figure 4.8.

(a) 3'-d(TTTTCTTTTCCCCCCT) T_m ~ 43 °C
5'-d(AGGGGGGA**X**AAGAAAA)

(b) 3'-d(TTTTCTTTTCCCCCCT) T_m ~ 12 °C or ~ 22 °C
5'-d(AGGGGGGA**X**AAGAAAA)

Figure 4.8 Proposed structure of (a) G.C-rich and (b) A.T-rich partial duplexes arising from possible mismatches between the TFO and target purine strand.

The higher T_m probably corresponds with the strongly paired, G.C-rich partial duplex (Figure 4.8 (a)) where X is not involved in binding and thus its structure would contribute little to the observed T_m. The lower T_m is likely to correspond to the more weakly paired, A.T-rich partial duplex (Figure 4.8 (b)) wherein the X base plays an active role in duplex formation. IsoG was accommodated in either its N3 or O2-H tautomeric form (Figure 4.9) in place of A without destabilisation of the partial duplex (T_m ~ 22 °C). However, in the

cases where X = B, L or G, duplex stability was hindered (T_m ~ 12 °C) due to a mismatch resulting from either electrostatic repulsion, steric hindrance or loss of H-bonding contacts.

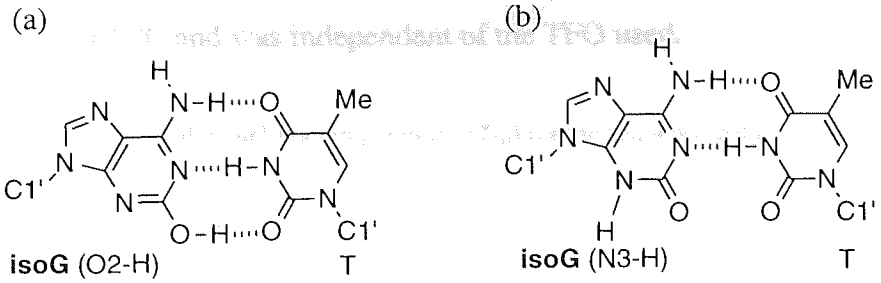


Figure 4.9 Base pairing of isoG with T via (a) the O2-H and (b) the N3-H tautomeric forms.

Horn et al[171] assessed the stability of a duplex containing an X.Y base pair (X = G or isoG, Y = C or T). The duplex containing the natural G.C base pair had T_m = 60 °C while the G.T and isoG.T containing duplexes had T_ms = 50 °C and 53 °C respectively. It was suggested that the O2-H tautomer of isoG was present at very low concentration under the conditions used (pH 7.9) since presumably a higher T_m would be expected for the isoG.T containing duplex based on the tridentate base pairing pattern shown in Figure 4.9 (a).

The results of this work have demonstrated the ability of isoG to form isoG.T base pairs which are equally as stable as A.T base pairs. As a result, it may be tentatively suggested that under the experimental conditions used here, isoG was present in its N3-H tautomeric form in the isoG.T base pair (Figure 4.9 (b)) where two H-bonds are present as in the isoenergetic A.T base pair. This would support the results of the molecular orbital calculations discussed in Section 4.1.3 where the N3-H tautomer was found to be the most thermodynamically favourable. Others[156, 172] have observed the pairing of isoG with T but suggested that the O2-H tautomer of isoG participated in the isoG.T base pair.

For the first time, isoG has been used to form triplexes which were found to be equally as stable as those formed by A, thereby extending the repertoire of bases which allow pH independent triplex formation.

4.7 Triplexation studies of sequences 242-245 and 247 at pH 6.4

4.7.1 Stability of TFOs containing isoG as a replacement for G in a TFO

The results of the thermal melting experiments carried out at pH 6.4 are outlined in Table 4.4. Spermine was added to the buffer solutions in all cases. For all cases studied at pH 6.4, the duplex had T_m ~ 61 °C and was independent of the TFO used.

Table 4.4 Duplex and triplex melting temperatures (T_m) for the sequence combinations listed.

Triplex	TFO	T_m (°C)	T_m (°C)	T_m (°C)
242*241.240	X = A, Y = G	33 ± 1	–	60.7 ± 0.5
247*241.240	X = A, Y = isoG	25 ± 1	54 ± 2	60.9 ± 0.5
244*241.240	X = L, Y = G	26 ± 1	41 ± 2	61.3 ± 0.5
245*241.240	X = B, Y = G	24 ± 1	54 ± 2	62.0 ± 0.5
243*241.240	X = isoG, Y = G	27 ± 1	40 ± 2	61.4 ± 0.5

TFO (**242**) contains four G-A or A-G junctions. Since the G*G.C triad is not isomorphous with the A*A.T triad, triplex **242*241.240** therefore contains four unfavourable steps in the triple strand. It was anticipated that the use of TFO (**247**), in which a single G is replaced with isoG, would lead to enhanced triplex stability since it contains only two unfavourable steps based upon the base-pairing pattern shown in Figure 4.1.

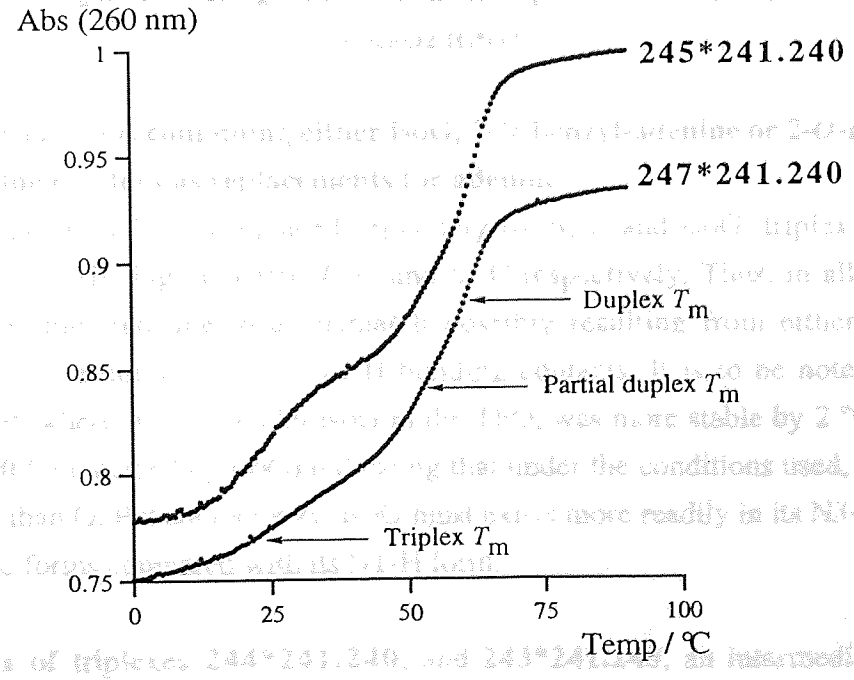


Figure 4.10 Melting curves corresponding to the triplexes **245*.241.240** and **247*241.240**.

The T_m of the unmodified triplex **242*241.240** was 33 ± 1 °C. When TFO (**247**) was used for triplexation, where G was replaced by isoG at position Y, the triplex had T_m = 25 ± 1 °C (Figure 4.10). Thus, isoG causes triplex destabilisation (~ 7 °C) when used in place of G. An intermediate transition with T_m = 54 ± 2 °C was also observed which is possibly associated with the formation of a partial duplex such as that shown in Figure 4.9 (a). Tautomerisation of the isoG base could result in mismatches between the N3-H or O2-H tautomeric forms of isoG and the G target of the duplex. Both of these tautomeric forms could only form a single H-bond with the target G forming highly unstable triads (Figure 4.11), ultimately resulting in triplex destabilisation. It is possible that the N1-H tautomeric form of isoG is not predominant in the microenvironment of the isoG*G.C triad.

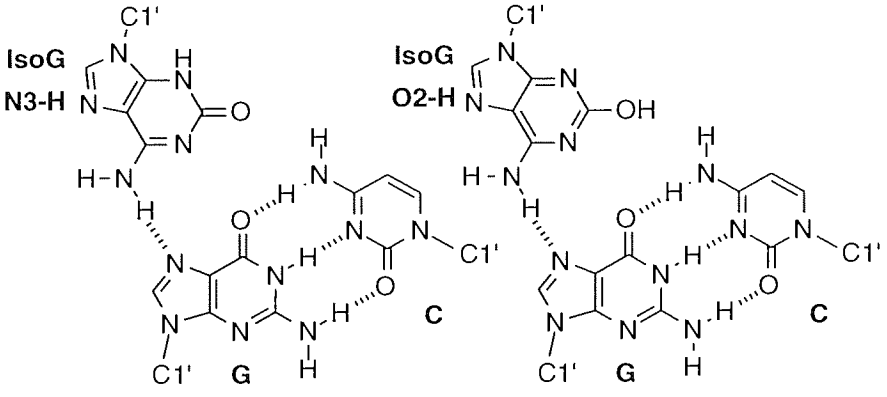


Figure 4.11 Pairing schemes for the base triplets isoG(N3-H)*G.C and isoG(O2-H)*G.C.

4.7.2 Study of TFOs containing either isoG, 2-*O*-benzyl-adenine or 2-*O*-allyl-adenine residues as replacements for adenine

When A in position X was replaced separately by B, L and isoG, triplex stability was reduced by 9 °C (B, Figure 4.10), 7 °C and 6 °C respectively. Thus, in all cases triplex stability was impaired due to a mismatch possibly resulting from either electrostatic repulsion, steric hindrance or loss of H-bonding contacts. It is to be noted that triplex **243*241.240**, where A replaced by isoG in the TFO, was more stable by 2 °C than triplex **247*241.240** (G replaced by isoG) indicating that under the conditions used, isoG behaves more like A than G. Put another way, isoG must exists more readily in its N3-H and/or O2-H tautomeric forms compared with its N1-H form.

In the cases of triplexes **244*241.240**, and **243*241.240**, an intermediate transition occurred at T_m ~ 41 °C which is similar to the intermediate transitions observed in the melting profiles of triplexes (**244-246**)***241.240** (Section 4.6, Table 4.2). However, in the

case of **245*241.240**, an intermediate transition was observed with T_m = 54 ± 2 °C similar to that observed for triplex **247*241.240**. Presuming that these intermediate transitions were due to the partial duplex shown in Figure 4.9 (a), the reason for the ~ 13 °C difference in T_m is unclear since the X base (B or L or isoG) is not involved in the partial duplex structure.

4.8 Conclusion

A new, direct synthetic route for the synthesis of 2'-deoxyisoguanosine (**217**) has been developed which is amenable to scale-up. It is likely that the yield of the hydrogenolysis reaction for the conversion of **227** to **217** could be improved significantly, possibly by more thorough washing of the Pd-C catalyst since it is possible that some of the product (**217**) adsorbed to the catalyst. The allyl-protected nucleoside **233** was incorporated into DNA sequences which allowed for generation of the isoG base within oligonucleotides via Pd(0)-mediated cleavage of the allyl group.

The studies carried out have demonstrated that isoG may be used as a replacement for A in a TFO without destabilisation of triple helices at both acidic and neutral pH, thus extending the repertoire of triplex-forming bases. It is thought that isoG was present in triplex **243*241.240** in either its N3-H or O2-H tautomeric form. The bases 2-*O*-allyl- and 2-*O*-benzyl-adenine were not accommodated well when used as replacements for adenine and caused marked triplex destabilisation due to a mismatch, possibly resulting from electrostatic repulsion, steric hindrance or loss of H-bonding contacts. A similar observation was made using TFO (**247**) where isoG was used as a replacement for G. In this case, isoG was surprisingly not accommodated in place of G suggesting that isoG was not present as its N1-H tautomer under the conditions used.

Future work in this area could involve stability studies of triplexes containing TFOs having more than one isoG residue targeted to a series of different DNA bases.

Chapter 5

Experimental

5.1 Chemistry

5.1.1 General Methods

NMR spectra were recorded on an a AC250 spectrometer at ^{1}H (250.1 MHz), ^{13}C (62.9 MHz), ^{19}F (235.3 MHz), ^{31}P (101.3 MHz). Positive chemical shifts are downfield of tetramethylsilane for ^{1}H and ^{13}C, downfield of 85% (aq) H_3PO_4 for ^{31}P and negative chemical shifts are upfield of CF_3CCl_3 for ^{19}F. Mass spectrometric analysis were carried out at either EPSRC (Swansea) in EI^+ or CI^+ mode using a VG Quatro II or FAB^+ mode using a VG AutoSpec instrument, or at the Department of Chemistry (Birmingham University) in EI^+ or CI^+ mode using a VG ProSpec or FAB^+ using a VG ZabSpec instruments or at Aston University in $APCI^+$ or ES (-) mode on a Hewlett-Packard HP 5989B MS Engine apparatus using a HP 59987A API-electrospray LC/MS interface. Infrared spectra were recorded on a Mattson Galaxy 2020 FT-IR Spectrophotometer. Ultraviolet spectra were recorded using Varian Cary E1 or Unicam PU8730 Spectrophotometers. Melting points were measured on Gallenkamp Electrothermal Digital apparatus and are uncorrected. Flash column chromatography was performed using Sorbsil C60 silica gel. TLC was performed using plastic-backed Kieselgel 60 silica gel plates containing a fluorescent indicator and visualised under UV (254 nm). Ethanolic anisaldehyde-H_2SO_4 was used where appropriate to develop TLC plates. Elemental Analyses were performed by Butterworth Laboratories (Middlesex).

5.1.2 Gifts

The metronidazole phosphoramidite (**182**) (Section 2.9) was a generous gift of Mike Davis, Department of Pharmaceutical Sciences, Aston University. The hydroxyhexyl derivatised CPG support (**99**) (Section 2.2.2) and the SLCPG support (**95**) were kind gifts of Dr. William Fraser, Department of Pharmaceutical Sciences, Aston University.

5.1.3 Synthesis

5-*N*-(4,4'-Dimethoxytrityl)-aminopentan-1-ol 107

5-Aminopentan-1-ol (**104**) (0.5 mL, 4.60 mmol) was dissolved in anhydrous pyridine (10 mL) and TMSCl (1.75 mL, 13.79 mmol) added dropwise at 0 °C under Ar and the mixture allowed to warm to r.t. A solution of DMTCl (1.55 g, 4.58 mmol) in anhydrous pyridine (5 mL) was then added and the clear solution was stirred at r.t. during 12 h. After this time MeOH (2 mL) was added and the solution stirred for 5 min before concentration of the

product mixture. The residue was dissolved in EtOAc (40 mL) and washed with ice-cold, saturated Na_2CO_3 (50 mL), brine (50 mL) and H_2O (50 mL) before drying over Na_2SO_4. Evaporation of the solvent gave a yellow oil which was purified by flash chromatography, eluting with EtOAc-hexane-Et_3N (40:60:1), affording the title compound **107** (1.06 g, 57%) as a viscous yellow oil; TLC [EtOAc-hexane-Et_3N (60:40:1)] R_f 0.56; IR (thin film): ν_{max} 3600-3100, 3075, 2929, 1604, 1511 cm^{-1}; ^1H NMR [$(CD_3)_2SO$]: δ 1.16, 1.4 (m, 6 H, $(CH_2)_3CH_2OH$), 1.94 (m, 2 H, CH_2NH), 2.37 (m, 1 H, NH), 3.34 (t, 2 H, *J* 6.3 Hz, CH_2OH), 3.7 (s, 6 H, 2 x CH_3O), 4.32 (t, 1 H, *J* 4.9 Hz, OH), 6.82 (d, 4 H *J* 8.3 Hz, 4 x CH (Ph)), 7.13-7.39 ppm (m, 9 H, 9 x CH (Ar)); ^{13}C NMR [$(CD_3)_2SO$]: δ 23.5, 30.1 ($CH_2CH_2CH_2NH$), 32.6 (CH_2CH_2OH), 43.5 (CH_2NH), 55.0 (2 x CH_3O), 60.7 (CH_2OH), 69.4 (Ph_3C), 113.0 (CH (Ph)), 125.8 (CH (Ph)), 127.6 (CH (Ph)), 128.2 (CH (Ph)), 129.5 (CH (Ph)), 138.6 (2 x C (Ph)), 147.0 (C (Ph)) and 157.3 ppm (2 x $COCH_3$). Mass spectrum (APCI$^+$): *m/z* (I_r) 406 (M + H, 1%), 303 (100%); Calcd for $C_{26}H_{31}NO_3$: (M + H) 406.238. Found 406.238.

5-[*N*-(4,4'-dimethoxytrityl)]aminopentyloxy-2-cyanoethoxy-*N*,*N*-diisopropylamino-phosphine 103

To alcohol (**107**) (0.28 g, 0.69 mmol) which was dried by co-evaporation with anhydrous pyridine (3 x 3 mL) and dissolved in anhydrous THF (4 mL), were added DIPEA (0.49 mL, 2.81 mmol) and 2-cyanoethyl-*N*,*N*-diisopropyl chlorophosphoramidite (0.23 mL, 1.03 mmol) with stirring under Ar. After 45 min, a precipitate formed, and the mixture concentrated to an oil which was purified by flash chromatography, eluting with EtOAc-hexane-Et_3N (25:75:1), affording the title compound **103** (0.37 g, 88%) as a clear yellow oil; TLC [EtOAc-hexane-Et_3N (25:75:1)]: R_f 0.27; IR (thin film): ν_{max} 3054, 2962, 1604, 1506, 1459 cm^{-1}; ^1H NMR [$(CD_3)_2SO$]: δ 1.08-1.13 (m, 12 H, 4 x CH_3 (Pr^i), 1.34-1.44 (m, 6 H, $(CH_2)_3CH_2NH$), 1.96 (m, 2 H, CH_2NH), 2.34 (t, 1 H, *J* 7.8 Hz, NH), 2.72 (t, 2 H, *J* 5.9 Hz, CH_2CN), 3.49-3.70 (m, 12 H, 2 x CH_3O, 2 x NCH, $POCH_2CH_2CN$, $POCH_2$), 6.8-7.39 ppm (m, 13 H, 13 x CH (Ar)); ^{13}C NMR [$(CD_3)_2SO$]: δ 19.8 (d, ^3JPC 6.9 Hz, CH_2CN), 23.4 ($(CH_2)_2CH_2NH$), 24.3 (2 x CH_3 (Pr^i)), 24.4 (2 x CH_3 (Pr^i)), 29.7 ($(CH_2)_2CH_2NH$), 30.7 (d, ^3JPC 7.1 Hz, $POCH_2CH_2CH_2$), 42.3 (NCH (Pr^i)), 42.5 (NCH (Pr^i)), 43.3 (CH_2NH), 54.9 (2 x CH_3), 58.0 (d, ^2JPC 18.3 Hz, $POCH_2CH_2CN$), 63.0 (d, ^2JPC 16.8 Hz, $POCH_2$), 69.3 (CNH), 112.9 (CH (Ar)), 119.0 (CN), 125.7 (CH (Ar)), 127.5 (CH (Ar)), 128.1 (CH (Ar)), 129.4 (CH (Ar)), 138.5 (CH (Ar)), 147.0 (C (Ar)), 157.2 ppm (2 x $COCH_3$); ^{31}P NMR ([$(CD_3)_2SO$]: δ 146.8 ppm; Mass spectrum (FAB$^+$): *m/z* (I_r) 628 (M + Na, 3%), 303 (100%); Calcd for $C_{35}H_{48}N_3O_4P$: (M + H) 606.346. Found 606.350.

3-(2-Chloroethyl)-4-oxoimidazo[5,1-*d*]-1,2,3,5-tetrazine-8-carboxylic acid 119

Mitozolomide (10.07 g, 0.04 mmol) was added slowly to concentrated sulphuric acid (50 mL) with stirring. Once dissolved, a solution of $NaNO_2$ (10.07 g, 0.15 mmol) in H_2O (25 mL) was added dropwise during 50 min, the temperature being kept below 60 °C. The resulting yellow solution was stirred for a further 3 h and poured onto ice giving a white precipitate which was then filtered off yielding the title compound **119** (9.59 g, 95%); mp 166 °C (lit.,[110] 166 °C); IR (KBr disc): ν_{max} 3122, 3000-2500, 1757 and 1710 cm^{-1}; ^1H NMR [$(CD_3)_2SO$]: δ 4.03 (t, 2 H, *J* 5.0 Hz, NCH_2), 4.65 (t, 2 H, *J* 5.0 Hz, CH_2Cl), 8.88 (s, 1 H, 6-CH), 13.5 ppm (br s, 1 H, CO_2H); Mass spectrum (FAB^+): *m/z* (I_r) 246 (M + H, 34%), 244 (M + H, 100%); Calcd for $C_7H_6ClN_5O_3$: C, 34.5%; H, 2.5%; N, 28.8%. Found C, 34.6%; H, 2.4%; N, 28.6%.

Similarly prepared was:

3-Methyl-4-oxoimidazo[5,1-*d*]-1,2,3,5-tetrazine-8-carboxylic acid 120

From temozolomide (2.01 g, 0.01 mmol) and $NaNO_2$ (2.60 g, 0.04 mmol) in H_2O (4 mL). The reaction temperature was kept at less than 38 °C. The title compound **120** (1.62 g, 80%) was isolated as a white solid; decomposes > 165 °C (lit.,[9] dec. 165-167 °C); ^1H NMR [$(CD_3)_2SO$]: δ (s, 3 H, NCH_3), 8.88 (s, 1 H, 6-CH), 13.5 ppm (br s, 1 H, CO_2H); ^{13}C NMR [$(CD_3)_2SO$]: δ 36.5 (CH_3), 127.8 (8-C or 8a-C), 129.3 (6-CH), 136.7 (8-C or 8a-C), 139.3 (4-C), 162.0 ppm (CO_2H); Mass spectrum (EI^+): *m/z* (I_r) 195 (M^+, 2%), 138 (15%), 57 (100%).

3-(2-Chloroethyl)-*O*-pentafluorophenyl-4-oxoimidazo[5,1-*d*]-1,2,3,5-tetrazine-8-carboxylate 136

The carboxylic acid **119** (1.23 g, 5.05 mmol) was dissolved in DMF (10 mL). C_6F_5OH (0.85 g, 4.62 mmol) in DMF (1.46 mL) and HOBT (0.68 g, 5.04 mmol) were added to the reaction mixture. The mixture was cooled to 5 °C before the addition of DCC (1.25 g, 6.05 mmol). After stirring for 3 h the solution was filtered through Celite and concentrated to a viscous brown oil. This was taken up in CH_2Cl_2 (5 mL) and filtered through Celite. Flash chromatography, eluting with hexane-EtOAc-MeCN (6:3:1) followed by precipitation into a solution of H_2O-MeOH (4:1), gave the title compound **136** (1.19 g, 63%) as a colourless solid; mp 126 °C; TLC [hexane-EtOAc-MeCN (6:3:1)]: R_f 0.43; IR (KBr disc): ν_{max} 3100, 1750, 1564 and 1515 cm^{-1}; ^1H NMR [$(CD_3)_2SO$]: δ 4.06 (t, 2 H, *J* 6.3 Hz, CH_2), 4.73 (t, 2 H, *J* 6.3 Hz, CH_2Cl), 9.11 ppm (s, 1 H, 6-CH); ^{19}F NMR [$(CD_3)_2SO$]: δ -15.55 (t, 2 F, *J* 25.0 Hz, 2 x *meta* CF), -20.38 (t, 1 F, *J* 23.8 Hz, *para* CF), -24.19 ppm (d, 2 F, *J* 20.0 Hz, 2 x *ortho* CF); Mass spectrum (FAB^+): *m/z* (I_r) 410 (M + H, 34%), 412 (M + H, 11%), 305

(100%), 226 (23%); Calcd for $C_{13}H_5ClF_5N_5O_3$: (M + H) 410.008. Found 410.009; Calcd for $C_{13}H_5ClF_5N_5O_3$: C, 38.1%; H, 1.2%, N, 17.1%. Found 38.5%; H, 1.2%; N, 16.9%.

Similarly prepared was:

3-Methyl-*O*-pentafluorophenyl-4-oxoimidazo[5,1-*d*]-1,2,3,5-tetrazine-8-carboxylate 137 from carboxylic acid **120** (0.16 g, 0.77 mmol), C_6F_5OH (0.16 g, 0.85 mmol) dissolved in DMF (0.27 mL), DCC (0.91 g, 0.93 mmol) and anhydrous DMF (7 mL) to give the title compound **137** as an off-white solid (0.08 g, 23%); mp 158 °C; TLC [hexane-EtOAc-MeCN (6:3:1)]: R_f 0.33; IR (KBr disc): ν_{max} 3118, 1755, 1560 and 1521 cm^{-1}; ^1H NMR [$(CD_3)_2SO$]: δ 3.95 (s, 3 H, CH_3), 9.07 ppm (s, 1 H, 6-CH); ^{19}F NMR [$(CD_3)_2SO$]: δ -15.60 (t, 2 F, *J* 20.4 Hz, 2 x *meta* CF), -20.40 (t, 1 F, *J* 23.3 Hz, *para* CF), -24.20 ppm (d, 2 F, *J* 19.6 Hz, 2 x *ortho* CF); Mass spectrum (FAB$^+$): *m/z* (I_r) 362 (M + H, 100%); Calcd for $C_{12}H_4F_5N_5O_3$: (M + H) 362.031. Found 362.032.

3-(2-Chloroethyl)-*N*-(5-hydroxypentyl)-4-oxoimidazo[5,1-*d*]-1,2,3,5-tetrazine-8-carboxamide 138

Method A: The active ester **136** (100 mg, 0.24 mmol) and 5-aminopentan-1-ol (24 μL, 0.22 mmol) were dissolved in a mixture of CH_2Cl_2 (3 mL) and MeOH (1.5 mL). The product solution was evaporated to dryness after 12 h and the residue subjected to flash chromatography, eluting with EtOAc-MeOH (9:1), to give the title compound **138** as a colourless solid (21.8 mg, 30%); mp 109 °C; TLC [EtOAc-MeOH (9:1)]: R_f 0.38; ^1H NMR [$(CD_3)_2SO$]: δ 1.31-1.57 (m, 6 H, (CH_2)$_3$$CH_2OH$), 3.29 (q, 2 H, *J* 6.6 Hz, CH_2NH), 3.37 (t, 2 H, *J* 6.5 Hz, CH_2OH (after D_2O exchange)), 4.03 (t, 2 H, *J* 6.0 Hz, CH_2CH_2Cl), 4.39 (t, 1 H, *J* 5.2 Hz, OH), 4.64 (t, 2 H, *J* 6.0 Hz, CH_2Cl), 8.55 ppm (t, 1 H, *J* 6.6 Hz, NH); ^{13}C NMR [$(CD_3)_2SO$]: δ 23.2 (CH_2), 29.3 (CH_2), 32.4 (CH_2CH_2OH), 38.8 (CH_2N), 41.7 (CH_2CH_2Cl), 50.1 (CH_2Cl), 60.8 (CH_2OH), 129.3 (6-CH), 131.5 (8-C or 8a-C), 134.0 (8-C or 8a-C), 139.3 (4-C), 159.6 ppm (C(O)NH). Mass spectrum (FAB$^+$): *m/z* (I_r) 329 (M + H, 100%), 331 (M + H, 33%); Calcd for $C_{12}H_{17}ClN_6O_3$: (M + H) 329.113. Found 329.113.

Method B: To 5-aminopentan-1-ol (0.19 g, 1.84 mmol), dissolved in DMF (10 mL), was added the carboxylic acid **119** (0.51 g, 2.09 mmol), PyBOP (1.09 g, 2.09 mmol) and DIPEA (1.00 mL, 5.75 mmol). The solution turned yellow immediately and then brown in colour. After 50 min the solvent was evaporated and the viscous black residue was purified by flash chromatography, eluting with EtOAc-MeOH (9:1), to give the title compound **138**

as a red solid (0.36 g, 60%). Analytical data were identical to those given above in Method A.

3-(2-Chloroethyl)-*N*-[5-*O*-(2-cyanoethyl-*N*,*N*-diisopropylphosphoramidite)pentyl]-4-oxoimidazo[5,1-*d*]-1,2,3,5-tetrazine-8-carboxamide 135

Alcohol **138** (0.39 g, 1.19 mmol) was co-evaporated with anhydrous THF (3 x 5 mL), dissolved in THF (10 mL) and placed under Ar. DIPEA (0.83 mL, 4.76 mmol) and 2-cyanoethyl *N*,*N*-diisopropylchlorophosphoramidite (0.32 mL, 1.40 mmol) were then added and the solution stirred for 2 h at r.t. after which time the solvent was evaporated and the residual brown oil was purified by flash column chromatography, eluting with EtOAc-CH_3CN-Et_3N (90:9:1), affording the title compound **135** (0.39 g, 62%) as a yellow oil; TLC [EtOAc-CH_3CN-Et_3N (90:9:1)]: R_f 0.68; IR (KBr disc): ν_{max} 3391, 2963, 2254, 1752, 1672 cm^{-1}; ^1H NMR [$(CD_3)_2SO$]: δ 1.12 (d, 6 H, *J* 5.0 Hz, 2 x CH_3), 1.13 (d, 6 H, *J* 5.0 Hz, 2 x CH_3), 1.40, 1.56 (2 x m, 6 H, (CH_2)$_2$$CH_2$NH)), 2.76 (t, 2 H, *J* 5.0 Hz, CH_2CN), 3.31 (m, 2 H, CH_2NH), 3.53-3.73 (m, 6 H, 2 x NCH (Pr^i_2), CH_2CH_2CN, $POCH_2$), 4.03 (t, 2 H, *J* 6.3 Hz, CH_2CH_2Cl), 4.64 (t, 2 H, *J* 6.3 Hz, CH_2Cl), 8.56 (m, 1 H, NH), 8.91 ppm (s, 1 H, 6-CH); ^{13}C NMR [$(CD_3)_2SO$]: δ 20.1 (d, ^3JPC 6.3 Hz, CH_2CN), 23.1 ($CH_2CH_2CH_2NH$), 24.5 (2 x CH_3), 24.6 (2 x CH_3), 29.0 ($CH_2CH_2CH_2NH$)), 30.6 (d, ^3JPC 7.6 Hz, $POCH_2CH_2$), 38.7 (CH_2NH), 41.7 (CH_2CH_2Cl), 42.5 (NCH (Pr^i)), 42.7 (NCH (Pr^i)), 50.1 (CH_2Cl), 58.3 (d, ^2JPC 18.2 Hz, CH_2CH_2CN), 63.1 (d, ^2JPC 19.6 Hz, $POCH_2$), 119.2 (CN), 129.3 (6-C), 131.5 (8-C), 134.0 (8a-C), 139.3 (4-C), 159.6 ppm (CONH); ^{31}P NMR [$(CD_3)_2SO$]: δ 146.7 ppm; Mass Spectrum (FAB$^+$): *m/z* (I_r) 553 (M + Na, 33%), 551 (M + Na, 100%); Calcd for $C_{21}H_{34}N_8O_4P$: (M + Na) 551.203. Found 551.203.

1-[2'-Deoxy-β-D-*erythro*-pentofuranosyl]-*N*-4-[2-trimethylsilylethoxycarbonyl]-pyrimidine-2*H*-one 146

To 2'-deoxycytidine (0.50 g, 2.04 mmol), dried by coevaporation with pyridine (3 x 10 mL) and dissolved in pyridine (10 mL), was added DMAP (0.25 g, 2.05 mmol) and the mixture was cooled to 0 °C under an Ar atmosphere. TMSCl (1.00 mL, 8.19 mmol) was added dropwise and the solution was stirred for 90 min after which a solution of 2-(trimethylsilyl)ethyl *p*-nitrophenyl carbonate (0.70 g, 2.47 mmol) in anhydrous pyridine (8 mL) was added and the reaction was heated to 60 °C for a further 12 h. After this time, H_2O (1 mL) and MeOH (1 mL) were added and the mixture was stirred at r.t. for further 3 h. The mixture was concentrated to an oil which was purified by flash chromatography, eluting with EtOAc-MeOH (9:1) affording the title compound **146** (0.49 g, 65%) as an off-white solid; mp 86-89 °C; TLC [MeOH-EtOAc (1:9)]: R_f 0.54; IR (KBr disc): ν_{max} 3394, 2956, 1745, 1653, 1500 cm^{-1}; ^1H NMR [$(CD_3)_2SO$]: δ 0.02 (s, 9 H, $(CH_3)_3$), 0.98 (t, 2 H *J*

8.4 Hz, CH_2Si), 2.02 (m, 1 H, 2'-CH_2), 2.23 (m, 1 H, 2'-CH_2), 3.57 (m, 2 H, 5'-CH_2), 3.83 (m, 1 H, 4'-CH), 4.14-4.21 (m, 3 H, $CH_2OC(O)$ and 3'-CH), 5.09 (s, 1 H, 3'-OH), 5.30 (s, 1 H, 3'-OH), 6.08 (t, 1 H *J* 6.21 Hz, 1'-CH), 7.03 (d, 1 H *J* 7.5 Hz, 5-CH), 8.29 (d, 1 H *J* 7.5 Hz, 6-CH), 10.67 ppm (br s, 1 H, NH); ^{13}C NMR [$(CD_3)_2SO$]: δ -1.2 ($(CH_3)Si$), 17.4 (CH_2Si), 41.1 (2'-CH_2), 61.2 (5'-CH_2), 63.7 (CH_2O), 70.2 (3'-CH), 86.3 (1'-CH), 88.1 (4'-CH), 94.5 (5-CH), 144.9 (6-CH), 153.6 (CO_2), 154.6 (4-C), 163.1 ppm (2-C); Mass spectrum (APCI$^+$): *m/z* (I_r) 372 (M + H, 3%), 228 (100%); Calcd for $C_{15}H_{25}N_3O_6Si$: C, 48.5%; H, 6.7%; N, 11.3%. Found C, 48.3%; H, 6.9%; N, 11.5%.

1-[2'-Deoxy-β-D-*erythro*-pentofuranosyl]-pyrimidine-2*H*-one 143

Nucleoside **146** (0.06 g, 0.16 mmol) was dissolved in a mixture of TBAF (0.5M) - AcOH (0.5M) in THF (0.42 mL). After 1h 10 min TLC analysis of the mixture indictated only starting material present [TLC: MeOH-EtOAc (1:9): R_f 0.30]. The mixture was then heated at 55 °C during 2 h. After this time TLC indicated that a new spot had appeared with complete disappearance of the starting material. This new spot corresponded to the title compound **143** as indicated by co-spotting (TLC) with an authentic sample of 2'-deoxycytosine; TLC [MeOH-EtOAc (1:9)]: R_f 0.07.

3-(2-Chloroethyl)-*N*-(5-iodopentyl)-4-oxoimidazo[5,1-*d*]-1,2,3,5-tetrazine-8-carboxamide 147

To alcohol **138** (0.57 g, 1.72 mmol) in MeCN (15 mL) cooled to 0 °C was added triphenylphosphine (0.55 g, 2.10 mmol) and DIPEA (0.6 mL, 3.44 mmol). After stirring for 15 min, I_2 (0.58 g, 2.29 mmol) was added and the reaction was stirred for a further 12 h after which time the reaction mixture was concentrated to give an oil which was dissolved in EtOAc (30 mL) and MeCN (10 mL) to give a clear solution. This solution was extracted with saturated $Na_2S_2O_3$ (20 mL). The organic layer was separated, and adsorbed onto silica. Flash chromatography, eluting with EtOAc-petroleum ether (bp 40-60 °C) (4:1) afforded the title compound **147** (0.23 g, 31% yield) as a brown solid; mp 74-77 °C; TLC [EtOAc-MeCN (4:1)]: R_f 0.77; ^1H NMR [$(CD_3)_2SO$]: δ 1.40, 1.59, 1.80 (3 x m, 3 x 2 H, $(CH_2)_3CH_2NH$), 3.26-3.31 (m, 4 H, CH_2I and CH_2NH), 4.03 (t, 2 H, *J* 6.0 Hz, CH_2CH_2Cl), 4.64 (t, 2 H, *J* 6.0 Hz, CH_2Cl), 8.60 (t, 1 H, *J* 5.8 Hz, NH), 8.92 ppm (s, 1 H, 6-CH); Mass spectrum (FAB$^+$): *m/z* (I_r) 439 (M + H, 100%), 334 (29%); Calcd for $C_{12}H_{16}ClIN_6O_2$: (M + H) 439.015. Found 439.015.

Attempted synthesis of 1-[2'-Deoxy-β-D-*erythro*-pentofuranosyl]-4-*N*-[5-[3-(2-Chloroethyl)-4-oxoimidazo[5,1-*d*]-1,2,3,5-tetrazine-8-carboxy]-aminopentyl]-pyrimidine-2*H*-one 151

To 2'-deoxycytidine (0.08 g, 0.34 mmol) dried by co-evaporation with anhydrous pyridine (3 x 5 mL), and suspended in pyridine (5 mL), was added TMSCl (0.10 mL, 0.79 mmol) and the solution stirred for 2 h, after which time a precipitate had formed and TLC analysis of the mixture indicated that silylation was complete. Iodo derivate **147** (0.12 g, 0.27 mmol) was added and the solution was stirred with heating at 50 °C during 24 h. After this time, TLC analysis only indicated the presence of the starting materials.

1-(2'-Deoxy-β-D-*erythro*-pentofuranosyl)-4-pentafluorophenyloxypyrimidin-2*H*-one 148

2'-Deoxyuridine (1.98 g, 8.67 mmol) was coevaporated with pyridine (3 x 5 mL), dissolved in pyridine (50 mL) and placed under Ar. Trifluoroacetic anhydride (3.5 mL, 24.78 mmol) was then added turning the solution a dark brown colour. After 63 h, C_6F_5OH (10.00 g, 54.33 mmol) dissolved in pyridine (10 mL) was added and the mixture was allowed to stir for a further 72 h. Concentration of the product mixture followed by flash chromatography of the resulting oil, eluting with EtOAc-MeOH 9:1, afforded the title compound **148** (2.47 g, 72%) as a light orange coloured solid; mp 176 °C (lit.,[125] 176 °C); TLC [EtOAc-MeOH (9:1)]: R_f 0.41; ^{1}H NMR [$(CD_3)_2SO$]: δ 2.08 (m, 1 H, 2'-CH), 2.29 (m, 1 H, 2'-CH), 3.62 (m, 2 H, 5'-CH_2), 3.88 (d, 1 H, *J* 3.5 Hz, 4'-CH), 4.23 (m, 1 H, 3'-CH), 5.15 (s, 1 H, 5'-OH), 5.29 (d, 1 H, *J* 4.2 Hz, 3'-OH), 6.06 (t, 1 H, *J* 6.1 Hz, 1'-CH), 6.60 (d, 1 H, *J* 7.3 Hz, 5-CH), 8.60 ppm (d, 1 H, *J* 7.3 Hz, 6-CH).

1-(2'-Deoxy-5'-*O*-(4,4'-dimethoxytrityl)-β-D-*erythro*-pentofuranosyl)-4-pentafluorophenylpyrimidin-2*H*-one 149

Nucleoside **148** (1.05 g, 2.67 mmol) was co-evaporated with anhydrous pyridine (3 x 5 mL) before dissolving in anhydrous pyridine (10 mL) under an Ar atmosphere to which DMTCl (1.12 g, 3.31 mmol) was added with stirring. After 90 min, MeOH (10 mL) was added to quench the reaction and the product mixture was evaorated to a black oil which was purified by flash chromatography, eluting with EtOAc-hexane-Et_3N (49:49:2), affording a cream-coloured solid which was confirmed as the title compound **149** when co-spotted with an authentic sample (1.19 g, 64% yield); mp 104 °C (lit.,[59] 104 °C).

1-[2'-Deoxy-5'-*O*-(4,4'-dimethoxytrityl)-β-D-*erythro*-pentofuranosyl]-4-*N*-(5-aminopentyl)-pyrimidine-2*H*-one 150

Nucleoside **149** (0.71 g, 1.02 mmol) was dissolved in anhydrous 1,4-dioxane (10 mL). 1,5-hexane diamine (0.60 mL, 5.10 mmol) was added and the mixture heated at 70 °C for 12 h during which time a precipitate formed. The mixture was evaporated to an oil and purified by flash chromatography, eluting first with MeOH-EtOAc-Et_3N (10:70:1) and then MeOH-Et_3N (90:10), affording the title compound **150** as a yellow foam (0.63 g, 85%); TLC [MeOH-Et_3N (90:10)]: R_f 0.10; ^{1}H NMR [$(CD_3)_2SO$]: δ 1.38 (m, 6 H, $(CH_2)_3CH_2NH_2$)), 1.90-2.25 (m, 2 H, 2'-CH_2), 3.17 (m, 6 H, 5'-CH_2, CH_2N, CH_2NH_2), 3.5-4.0 (m, 10 H, NH_2, 2 x CH_3O, 3'-OH, 4'-CH), 4.24 (m, 1 H, 3'-CH), 5.59 (d, 1 H *J* 7.4 Hz, 5-CH), 6.18 (t, 1 H *J* 6.7 Hz, 1'-CH), 6.88-7.37 (m, 13 H, 13 x CH (Ph)), 7.59 (d, 1 H *J* 7.4 Hz, 6-CH), 7.75 ppm (m, 1 H, NH); ^{13}C NMR [$(CD_3)_2SO$]: δ 23.4 (CH_2), 27.7 (CH_2), 28.1 (CH_2), 38.5 (CH_2N), 39.0 (CH_2NH_2), 45.7 (2'-CH_2), 55.1 (2 x CH_3O), 66.4 (5'-CH_2), 70.1 (3'-C), 84.6 (1'-CH), 85.1 (4'-CH), 85.8 (Ar_3C), 94.7 (5-CH), 113.3 (CH (Ar)), 127.7 (CH (Ar)), 127.9 (CH (Ar)), 129.8 (CH (Ar)), 135.3 (C), 135.5 (C), 139.4 (6-CH), 144.7 (PhC), 155.1 (4-C), 158.1 (2 x $COCH_3$) and 163.2 ppm (2-C); Mass spectrum (FAB$^+$): *m/z* (I_r) 637 (M + Na, 16%), 303 (100%); Calcd for $C_{35}H_{42}N_4O_6$: (M + H) 615.318. Found 615.317.

1-[2'-Deoxy-5'-*O*-(4,4'-dimethoxytrityl)-β-D-*erythro*-pentofuranosyl]-4-*N*-[5-[3-(2-Chloroethyl)-4-oxoimidazo[5,1-*d*]-1,2,3,5-tetrazine-8-carboxy]-aminopentyl]-pyrimidine-2*H*-one 151

The ester (**136**) (0.05 g, 0.08 mmol) was dissolved in DMSO (0.5 mL) to which was added nucleoside **150** (0.04 g, 0.10 mmol) and the mixture was stirred for 1 h. The product was purified by flash chromatography eluting firstly with 10% and then 20% MeOH-EtOAc affording the title compound **151** (20 mg, 31%) as a cream-coloured solid; TLC [EtOAc-MeOH (9:1)]: R_f 0.43; ^{1}H NMR [$(CD_3)_2SO$]: δ 1.20-1.60 (m, 6 H, $(CH_2)_3CH_2NH$), 2.05 (m, 1 H, 2'-CH), 2.15 (m, 1 H, 2'-CH), 3.15-3.35 (m, 4 H, CH_2NH, 5'-CH_2), 3.74 (s, 6 H, 2 x CH_3), 3.85 (m, 1 H, 4'-CH), 4.02-4.10 (m, 5 H, CH_2CH_2Cl, $CH_2NHC(O)$, 3'-OH), 4.25 (m, 1 H, 3'-CH), 4.63 (t, 2 H, *J* 5.0 Hz, CH_2Cl), 5.58 (d, 1 H, *J* 7.4 Hz, 5-CH), 6.17 (t, 1 H, *J* 6.1 Hz, 1'-CH), 6.90 (m, 4 H, 4 x CH (Ar)), 7.15-7.40 (m, 9 H, 9 x CH (Ar)), 7.59 (d, 1 H, *J* 7.6 Hz, 6-CH), 7.71 (m, 1 H, NH), 8.58 (m, 1 H, NH), 8.90 ppm (s, 1 H, CH (imidazole)).

1-[2'-Deoxy-5'-*O*-(4,4'-dimethoxytrityl)-β-D-*erythro*-pentofuranosyl]-4-*N*-[5-(*N*-allyloxycarbonyl)aminopentyl]-pyrimidine-2*H*-one 156

Nucleoside **150** (1.01 g, 1.65 mmol) was co-evaporated with anhydrous pyridine (3 x 10 mL), dissolved in pyridine (50 mL), placed under Ar and the solution cooled to 0 °C. TMSCl (0.42 mL, 3.30 mmol) was added dropwise and the solution stirred for 40 min. Allyloxycarbonyl chloroformate (0.25 mL, 2.13 mmol) was then added and the mixture was stirred for 17 h after which time H_2O (0.4 mL) and MeOH (0.4 mL) were added and the mixture stirred at r.t. for a further 3 h. The mixture was evaporated to an oil and purified by flash chromatography eluting firstly with EtOAc-Et_3N (99:1) and then with MeOH-EtOAc-Et_3N (10:89:1), to afford the title compound **156** (0.53 g, 46%) as a yellow foam; mp 93-96 °C; TLC [MeOH-EtOAc-Et_3N (10:89:1)]: R_f 0.4; IR (KBr disc): ν_{max} 3312, 2937, 1701, 1641, 1512 cm^{-1}; ^1H NMR [$(CD_3)_2SO$]: δ 1.31-1.46 (m, 6 H, $(CH_2)_3CH_2NH$), 2.07-2.20 (m, 2 H, 2'-CH_2), 2.98 (m, 2 H , CH_2NH), 3.18-3.30 (m, 4 H, 5'-CH_2, CH_2NH), 3.75 (s, 6 H, 2 x CH_3), 3.87 (m, 1 H, 4'-CH), 4.26 (m, 1 H, 3'-CH), 4.46 (d, 2 H *J* 5.3 Hz, OCH_2 (allyl)), 5.15-5.31 (m, 3 H, 3'-OH and CH_2 (allyl)), 5.60 (d, 1 H *J* 7.4 Hz, 5-CH), 5.87 (m, 1 H, CH (allyl)), 6.18 (t, 1 H *J* 6.1 Hz, 1'-CH), 6.89 -7.41 (m, 14 H, 13 x CH (Ar), NH), 7.62 (d, 2 H *J* 7.6 Hz, 6-CH), 7.81 ppm (m, 1 H, NH); ^{13}C NMR [$(CD_3)_2SO$]: δ 23.8 ($(CH_2)_3CH_2NH$), 28.2 ($(CH_2)_3CH_2NH$), 29.2 ($(CH_2)_3CH_2NH$), 39.9, 40.2, 40.4 (2'-CH, 2 x CH_2NH), 55.1 (2 x CH_3O (DMT)), 63.4 (5'-CH_2), 64.1 (CH_2O (allyl)), 70.0 (3'-CH), 84.6 (1'-CH), 85.2 (4'-CH), 85.8 (Ar_3C), 94.6 (5-CH), 113.3 (CH (Ar)), 116.9 (CH_2 (allyl)), 126.8 (CH (Ar)), 127.7 (CH (Ar)), 127.9 (CH (Ar)), 129.8 (CH (Ar)), 133.9 (CH (allyl)), 135.4 (C (Ar)), 135.5 (C (Ar)), 139.5 (6-CH), 144.8 (C (Ar)), 154.7 (4-C), 156.0 (CO_2 (allyl)), 158.1 (2 x $COCH_3$ (Ar)), 163 ppm (2-C); Mass spectrum (FAB^+): *m/z* (I_r) 721 (M + Na, 8%), 303 (100%); Calcd for $C_{39}H_{46}N_4O_8$: (M + Na) 721.321. Found 721.322. Calcd for $C_{39}H_{46}N_4O_8$: C, 67.0%; H, 6.6%; N, 8.0%. Found C, 66.5%; H, 6.6%; N, 8.1%.

1-[2'-Deoxy-5'-*O*-(4,4'-dimethoxytrityl)-β-D-*erythro*-pentofuranosyl]-4-*N*-[5-(*N*-allyloxycarbonyl)aminopentyl]-pyrimidine-2*H*-one-3'-*O*-(2-cyanoethyl)-*N*,*N*-diisopropylphosphoramidite 157

Nucleoside **156** (0.36 g, 0.52 mmol) was coevaporated with anhydrous pyridine (3 x 5 mL) and dissolved in anhydrous THF (7 mL) and placed under Ar. DIPEA (0.36 mL, 2.07 mmol) and 2-cyanoethyl *N*,*N*-diisopropylchlorophosphoramidite (0.24 mL, 1.08 mmol) were added and the mixture stirred for 105 min after which time the mixture was concentrated to an oil and the product purified by flash chroamtography eluting with EtOAc-MeCN (9:1) affording the title compound **157** (0.27 g, 39%) as a colourless viscous

oil which was used directly for synthesis of sequence **158** (Section 5.3.20); TLC [EtOAc-MeCN (9:1)]: R_f 0.63; ^{31}P NMR [$(CD_3)_2SO$]: δ 147.8 (34.5%), 147.6 ppm (34.5%).

t-Butyl-2-hydroxyethylamine-N-carboxylate 160

2-Aminoethanol (1.01 g, 16.6 mmol) was dissolved in 1,4-dioxane (15 mL) and Boc-ON (3.65 g, 14.82 mmol) was added. The solution was stirred for 3 h after which time the solvent was evaporated and the residual viscous yellow oil was purified by flash chromatography, eluting with hexane-EtOAc-MeOH (7:2:1) to give the title compound **160** as a yellow oil (1.79 g, 75%); TLC [EtOAc-MeOH (9:1)]: R_f 0.49; IR (KBr disc): ν_{max} 3352, 2975 and 1693 cm^{-1}; ^{1}H NMR [$(CD_3)_2SO$]: δ 1.39 (s, 9 H, Bu^t), 2.98 (q, 2 H, J 5.0 Hz, CH_2NH), 3.33 (t, 2 H, J 6.3 Hz, CH_2OH), 4.60 (br s, 1 H, OH), 6.75 ppm (m, 1 H, NH); ^{13}C NMR [$(CD_3)_2SO$]: δ 28.5 (Bu^t), 42.9 (CH_2NH), 60.3 (CH_2OH), 77.7 (Bu^{t}C), 155.9 ppm (CO); Mass spectrum (CI$^+$): m/z (I_r) 162 (M + H, 10%), 105 (100%); Calcd for $C_7H_{15}NO_3$: C, 52.2%; H, 9.3%; N, 8.7%. Found: C, 51.9%; H, 9.2%; N, 8.8%.

t-Butyl-2-hydroxypentylamine-N-carboxylate 162

5-Aminopentan-1-ol (2.00 g, 19.39 mmol) dissolved in 1,4-dioxane (20 mL) was added Boc-ON (4.56 g, 18.50 mmol) and the solution was stirred for 3 h after which time the solvent was evaporated to give a yellow viscous oil which was purified by flash chromatography, eluting with hexane-EtOAc-MeOH (7:2:1) afforded the title compound **162** (2.95 g, 79%) as a yellow oil; TLC [hexane-EtOAc-MeOH (7:2:1)] R_f 0.13; IR (thin film): ν_{max} 3354, 2937 and 1687 cm^{-1}; ^{1}H NMR [$(CD_3)_2SO$]: δ 1.25-1.45 (m, 15 H, Bu^t, (CH_2)$_3$CH$_2$OH), 2.89 (q, 2 H, J 7.5 Hz, CH_2NH), 3.37 (q, 2 H, J 7.5 Hz, CH_2OH), 4.35 (br s, 1 H, OH), 6.75 ppm (m, 1 H, NH); ^{13}C NMR [$(CD_3)_2SO$]: δ 23.3 (CH_2), 28.7 (Bu^t), 29.9 (CH_2), 32.7 (CH_2), 40.4 (CH_2NH), 61.1 CH_2OH), 77.7 (Bu^{t}C), 156.0 (CO); Mass spectrum (APCI$^+$): m/z (I_r) 204 (M + H, 13%); Calcd for $C_{10}H_{21}NO_3$: (M + H) 204.156. Found 204.160.

3-(2-Chloroethyl)-N-(2-(ammoniumtrifluoroacetyl)ethyl)-4-oxoimidazo[5,1-d]-1,2,3,5-tetrazine-8-carboxylate 164

To alcohol **160** (0.27 g, 1.67 mmol) in DMF (3 mL) at r.t. was added **119** (0.45 g, 1.84 mmol), PyBOP (0.87 g, 1.68 mmol) and DIPEA (0.8 mL, 4.60 mmol). Initially the solution turned yellow in colour and then yellow/brown in colour. After 45 min the solvent was evaporated afording a black viscous residue which was purified by flash chromatography, eluting with Et$_2$O-EtOAc-MeCN (6:3:1), affording 3-(2-chloroethyl)-N-(2-t-butoxycarbonylaminoethyl)-4-oxoimidazo[5,1-d]-1,2,3,5-tetrazine-8-carboxylate **163** (yield not recorded, TLC [EtOAc-MeOH (9:1)] R_f 0.64). The solid was dissolved in a 90%

TFA-CH_2Cl_2 solution (2 mL) and stirred during 1 h after which time the product mixture was diluted with CH_2Cl_2 (12 mL). The organic phase was extracted with H_2O (2 x 15 mL). The combined aqueous layers were re-extracted with CH_2Cl_2 (20 mL). The aqueous phase was then evaporated affording the title compound **164** (0.49 g, 73% over 2 steps) as a brown viscous oil; ^{1}H NMR [$(CD_3)_2SO$]: δ 3.25 (m, 2 H, $CH_2NH_3^+$), 4.05 (m, 2 H, CH_2CH_2Cl), 4.57 (m, 2 H, CH_2O), 4.63 (m, 2 H, CH_2Cl), 8.05 (br s, 3 H, NH_3^+), 8.98 ppm (s, 1 H, 6-CH); ^{13}C NMR [$(CD_3)_2SO$]: δ 30.9 (CF_3), 38.1 ($CH_2NH_3^+$), 41.8 (CH_2CH_2Cl), 50.7 (CH_2Cl), 61.7 (CH_2O), 127.0 (6-C), 130.4 (8-C or 8a-C), 136.6 (8-C or 8a-C), 139.0 (4-CO), 160.3 (CO_2), 161.8 ppm (CO_2^-); ^{19}F NMR [$(CD_3)_2SO$]: δ 6.02 ppm; Mass spectrum (FAB$^+$): (I_r) 287 (M - $CF_3CO_2^-$, 100%), 289 (M - $CF_3CO_2^-$, 33%); Calcd for $C_7H_{15}NO_3$: 287.066. Found 287.066.

3-(2-Chloroethyl)-*N*-(5-*t*-butoxycarbonylaminopentyl)-4-oxoimidazo[5,1-*d*]-1,2,3,5-tetrazine-8-carboxylate 165

To alcohol **162** (0.41 g, 2.02 mmol) in DMF (3 mL) was added **119** (0.53 g, 2.18 mmol), PyBOP (1.03 g, 1.98 mmol) and DIPEA (0.95 mL, 5.45 mmol) at r.t.. Initially the solution turned yellow in colour and then yellow/brown in colour. After 45 min the solvent was evaporated affording a black viscous residue which was purified by flash chromatography, eluting with Et_2O-EtOAc-MeCN (6:3:1), affording the title compound **165** (0.59 g, 70%), mp 106 °C; TLC [Et_2O-EtOAc (7:3)]: R_f 0.49; IR (KBr disc): ν_{max} 3365, 3114, 2932, 1735, 1715 and 1683 cm^{-1}; ^{1}H NMR [$(CD_3)_2SO$]: δ 1.30-1.43 (m, 15 H, Bu^t and $(CH_2)_3CH_2NH$), 2.90-2.93 (m, 2 H, CH_2NH), 4.03 (t, 2 H, *J* 6.0 Hz CH_2CH_2Cl), 4.34 (t, 2 H, *J* 6.5 Hz, CH_2O), 4.67 (t, 2 H, *J* 6.0 Hz, CH_2Cl), 6.82 (br s, 1 H, NH), 8.92 ppm (6-CH); ^{13}C NMR [$(CD_3)_2SO$]: δ 22.8 $(CH_2)_3CH_2OH$), 28.0 $(CH_2)_3CH_2OH$), 28.3 (Bu^t), 29.2 $(CH_2)_3CH_2OH$), 39.8 (CH_2NH), 41.5 (CH_2CH_2Cl), 50.3 (CH_2Cl), 64.9 (CH_2O), 77.4 (Bu^{t}C), 127.4 (8-C or 8a-C), 130.1 (6-CH), 136.2 (8-C or 8aC), 138.9 (4-C), 155.7 (OC(O)NH), 160.5 ppm (CO_2); Mass spectrum (FAB$^+$): *m/z* (I_r) 429 (M + H, 81%), 431 (M + H, 27%), 329 (100%). Calcd for $C_{17}H_{25}ClN_6O_5$: (M + H) 429.165. Found 429.166.

3-(2-Chloroethyl)-*N*-(5-(ammoniumtrifluoroacetyl)pentyl)-4-oxoimidazo[5,1-*d*]-1,2,3,5-tetrazine-8-carboxylate 166

165 (0.11 g, 0.26 mmol) was dissolved in 90% TFA-CH_2Cl_2 solution (5 mL) and stirred during 1 h after which time the product mixture was diluted with CH_2Cl_2 (12 mL). The organic phase was extracted with H_2O (2 x 15 mL). The combined aqueous layers were re-extracted with CH_2Cl_2 (20 mL). The aqueous phase was then evaporated affording the title compound **166** as a viscous oil (0.11 g, 99%); ^{1}H NMR [$(CD_3)_2SO$]:δ 1.40-1.80 (m, 6 H, $(CH_2)_3CH_2NH_3^+$), 2.79 (m, 2 H, $CH_2NH_3^+$), 4.04 (t, 2 H, *J* 6.0 Hz, CH_2CH_2Cl), 4.37 (t, 2

H, *J* 6.4 Hz, CH_2O), 4.67 (t, 2 H, *J* 6.0 Hz, CH_2Cl), 7.75 (br s, 3 H, NH_3^+), 8.94 ppm (s, 1 H, 6-CH); ^{13}C NMR [$(CD_3)_2SO$]: δ 22.3 ($(CH_2)_3CH_2O$), 26.6 ($(CH_2)_3CH_2O$), 27.6 ($(CH_2)_3CH_2O$), 34.3 (CF_3), 38.7 ($CH_2NH_3^+$), 41.5 (CH_2CH_2Cl), 50.3 (CH_2Cl), 64.6 (CH_2O), 127.2 (8-C or 8a-C), 130.1 (6-CH), 136.1 (8a-C or 8-C), 138.9 (4-C), 160.4 (CO_2), 192.7 ppm (CO_2^-); ^{19}F NMR [$(CD_3)_2SO$]: δ 6.03 ppm; Mass spectrum (FAB^+): *m/z* (I_r) 329 (M - $CF_3CO_2^-$, 100%), 331 (M - $CF_3CO_2^-$, 33%); Calcd for $C_{12}H_{18}ClN_6O_3$: (M - $CF_3CO_2^-$) 329.113. Found 329.114.

Attempted syntheses of 1-[2'-Deoxy-β-D-*erythro*-pentofuranosyl]-4-*N*-[5-[3-(2-Chloroethyl)-4-oxoimidazo[5,1-*d*]-1,2,3,5-tetrazine-8-carbonyl]-aminoethyl]-pyrimidine-2*H*-one 167

Method A: To **164** (0.13 g, 0.33 mmol), dissolved in anhydrous 1,4-dioxane (6 mL) under Ar, was added nucleoside **148** (0.13 g, 0.33 mmol) and Et_3N (0.14 mL, 1.00 mmol) and the mixture was stirred with heating at 70 °C. After 3 h, no evidence of reaction was observed by TLC and so DMAP (0.06 g, 0.52 mmol) was added and the reaction was continued for another 12 h. After this time, TLC indicated only the presence of starting materials.

Method B: To **164** (9.3 mg, 23 μmol) dissolved in anhydrous 1,4-dioxane (1 mL) under Ar was added nucleoside **148** (9.8 mg, 25 μmol) and DIPEA (12 μL, 69 μmol) and the mixture was stirred with heating at 70 °C during 12 h. After this time, TLC indicated only the presence of starting materials.

Attempted synthesis of 1-[2'-Deoxy-β-D-*erythro*-pentofuranosyl]-4-*N*-[5-[3-(2-Chloroethyl)-4-oxoimidazo[5,1-*d*]-1,2,3,5-tetrazine-8-carbonyl]-aminopentyl]-pyrimidine-2*H*-one 168

To **166** (0.10 mg, 0.22 mmol), dissolved in anhydrous 1,4-dioxane (5 mL) under Ar, was added nucleoside **148** (0.11 g, 0.28 mmol) and DIPEA (0.19 mL, 1.10 mmol) and the mixture was stirred with heating at 70 °C during 12 h. After this time, TLC indicated only the presence of starting materials.

3-(2-Chloroethyl)-*N*-*n*-butyl-4-oxoimidazo[5,1-*d*]-1,2,3,5-tetrazine-8-carboxamide 169

To the active ester **136** (0.16 g, 0.39 mmol) dissolved in CH_2Cl_2 (10 mL) was added Bu^nNH_2 (0.1 mL, 1.01 mmol) with stirring and the solution turned yellow. The precipitated brown title compound (**169**) was filtered off after 45 min (70.3 mg, 60%); TLC [MeOH-EtOAc(1:1)] : R_f 0.09; ^1H NMR [$(CD_3)_2SO$]: δ 0.90 (t, 3 H, *J* 7.3 Hz, CH_3CH_2), 1.32 (sextet, 2 H, *J* 7.4 Hz, CH_3CH_2), 1.50 (pentet, 2 H, *J* 7.2 Hz, CH_2CH_2NH), 3.26-3.33 (q, 2

H, *J* 6.3 Hz, CH_2N), 4.03 (t, 2 H, *J* 6.0 Hz, CH_2CH_2Cl), 4.64 (t, 2 H, *J* 6.0 Hz, CH_2Cl), 8.55 (t, 1 H, *J* 6.0 Hz, NH), 8.91 ppm (s, 1 H, 6-CH).

Similarly prepared was:

3-Methyl-*N*-*n*-butyl-4-oxoimidazo[5,1-*d*]-1,2,3,5-tetrazine-8-carboxamide 171

From the ester **137** (0.06 g, 0.17 mmol) and Bu^nNH_2 (15.2 µl, 0.15 mmol) in CH_2Cl_2 (5 mL) to give the title compound (**171**) as an orange/brown solid (0.03 g, 58%); mp 151 °C; TLC [toluene-EtOAc (85:15)]: R_f 0.10; IR (KBr disc): ν_{max} 3317, 3081, 2930, 1742 and 1652 cm^{-1}; ^1H NMR [$(CD_3)_2SO$]: δ 0.90 (t, 3 H, *J* 7.4 Hz, CH_3CH_2), 1.33 (sextet, 2 H, *J* 7.5 Hz, CH_3CH_2), 1.52 (pentet, 2 H, *J* 7.6 Hz, CH_2CH_2NH), 3.29 (q, 2 H, *J* 7.5 Hz, CH_2NH), 3.87 (s, 3 H, CH_3N), 8.49 (t, 1 H, *J* 6.3 Hz, NH), 8.85 ppm (s, 1 H, 6-CH); Mass spectrum (FAB^+): *m/z* (I_r) 251 (M + H, 100%); Calcd for $C_{10}H_{14}N_6O_2$: 251.126. Found 251.126.

General procedure for attempted preparation of 3-(2-Chloroethyl)-*N*-(2-aminoethyl)-4-oxoimidazo[5,1-d]-1,2,3,5-tetrazine-8-carboxamide 170 (n = 1) and 3-(2-Chloroethyl)-*N*-(5-aminopropyl)-4-oxoimidazo[5,1-d]-1,2,3,5-tetrazine-8-carboxamide 170 (n = 2)

Diamine (ethylene diamine or 1, 3-diaminopropane) (5.67 mmol) was dissloved in CH_2Cl_2 (5 mL). The active ester **136** (0.14 g, 0.35 mmol) was dissolved in CH_2Cl_2 (2.5 mL) and added dropwise, during 5 min, to the solution of the diamine. The precipitate which formed was isolated by filtration after 45 min. ^1H NMR showed that the isolated products were neither the starting materials nor the expected title compounds. The identity of the compounds in the compound mixture could not be established.

***t*-Butyl 2-benzylaminoethylamine-*N*-carboxylate 173**

Boc-ON (0.89 g, 3.59 mmol) was added to a solution of benzylethylenediamine (**172**) (0.49 g, 3.26 mmol) in 1,4-dioxane (8 mL) and the reaction was stirred for 3 h. The solvent was evaporated leaving a yellow oil which was purified by flash chromatography using firstly, hexane-EtOAc-MeCN (7:2:1) to elute the oxime side-product and, then EtOAc-MeOH (9:1) to elute the title compound **173** as a colourless solid (0.60 g, 74%); TLC [EtOAc-MeOH (4:1)] R_f 0.28; ^1H NMR [$(CD_3)_2SO$]: δ 1.38 (s, 9 H, Bu^t), 2.53 (t, 2 H, *J* 6.3Hz), 2.75 (br s, 1 H, NHBoc), 3.04 (q, 2 H, *J* 6.1 Hz, CH_2NH), 3.67 (s, 2 H, CH_2Ph), 6.8 (br s, 1 H, NHBoc), 7.21-7.32 ppm (m, 5 H, C_6H_5); ^{13}C NMR [$(CD_3)_2SO$]: δ 28.4 (CH_3), 40.2 (CH_2Ph), 48.6 (CH_2N), 52.9 (CH_2NHBoc), 77.7 (Bu^tC), 126.7 (CH), 128.1 (CH), 128.3 (CH), 141.1 (C), 155.9 ppm (CO).

Attempted synthesis of *t*-Butyl 2-aminoethylamine-*N*-carboxylate 174

t-Butyl 2-benzylaminoethylamine-*N*-carboxylate **173** (0.60 g, 2.40 mmol) was dissolved in MeOH (30 mL) to which was added 10% Pd on charcoal (60 mg). The mixture was stirred and placed under a hydrogen atmosphere. After 7 days, the reaction mixture was filtered through Celite and the solvent evaporated to give an unidentified viscous yellow oil (0.29 g, 75%); TLC [EtOAc-MeOH (4:1)]: R_f 0.17; ^{1}H NMR [$(CD_3)_2SO$]: δ 0.92-0.95 (d, 6 H, *J* 6.2 Hz, identity unknown; singlet upon decoupling at 2.66 ppm), 1.37 (s, 9 H, Bu^t), 2.46-2.52 (t, 2 H, *J* 7.5 Hz, CH_2NH$_2$), 2.66 (septet, 1 H, *J* 6.3 Hz, identity unknown), 2.97-2.99 (m, 2 H, CH_2NH), 3.2 (s, 2 H, PhCH_2), 6.78 ppm (br s, 1 H, NH); Mass spectrum (APCI$^+$): *m/z* (I_r) 203 (unidentified, 55%), 147 (unidentified, 64%), 129 (unidentified, 75%). Expected 161 (M + H).

***t*-Butyl-2-cyanoethylamine-*N*-carboxylate 176**

To a solution of aminopropionitrile (0.52 g, 7.43 mmol) in 1,4-dioxane (5 mL) stirred at 5 °C was added Boc-ON (1.83 g, 7.43 mmol) giving a clear, colourless solution. The solution was stirred for 3 h before the solvent was evaporated and the residue was dissolved in CH_2Cl_2 (20 mL). The organic phase was extracted with saturated Na_2CO_3 solution (2 x 20 mL), H_2O (1 x 20 mL) and the solvent evaporated to give the title compound **176** (0.84 g, 67%) as a brown solid; mp 42 °C; TLC [EtOAc-MeOH (1:1)] R_f 0.12; IR (thin film): ν_{max} 3363, 2987, 2244 and 1687 cm^{-1}; ^{1}H NMR [$(CD_3)_2SO$]: δ 1.39 (s, 9 H, Bu^t), 2.60 (t, 2 H, *J* 6.3 Hz, CH_2CN), 3.16 (q, 2 H, *J* 7.5 Hz, CH_2N), 7.18 (br s, 1 H, NH); Mass spectrum (APCI$^+$): *m/z* (I_r) 171 (M + H, 93%); Calcd for $C_8H_{14}N_2O_2$: C, 56.5%; H, 8.2%; N, 16.5%. Found: C, 56.4%; H, 8.0%; N, 16.5%.

Attempted synthesis of *t*-Butyl-2-aminopropylamine-*N*-carboxylate 178

Method A: To nitrile (**176**) (0.41 g, 2.41 mmol) dissolved in MeOH (35 mL) was added 10% Pd-C (0.17 g) and the mixture was hydrogenolised at r.t. for 6 days using a hydrogen baloon. After this time TLC indicated that 5 components were present. No purification was therefore attempted.

Method B: To nitrile (**176**) (0.20 g, 1.18 mmol) dissolved in a 10% MeOH-NH_3 solution (30 mL) was added 5% Rh on C (65mg) and the mixture was hydrogenolysed under an atmosphere of hydrogen in an hydrogenator at 80 psi. After 60 h TLC indictated that 8 components were present. No purification was therefore attempted.

1-(3',5'-Di-*O*-acetyl-2'-deoxy-β-D-*erythro*-pentofuranosyl)-5-methyl-pyrimidine-2,4-(1*H*,3*H*)-dione 191

To 2'-deoxythymidine (6.62 g, 27.35 mmol), dried by coevaporation with anhydrous pyridine (3 x 30 mL) and dissolved in anhydrous pyridine (100 mL) at 0 °C, was added Ac_2O (26.0 mL, 275.56 mmol). After 18 h the reaction mixture was evaporated to give the title compound **191** as a white foam (yield not recorded); TLC (EtOAc): R_f 0.35 (lit.,[130] 0.34)).

1-(3',5'-Di-*O*-acetyl-2'-deoxy-β-D-*erythro*-pentofuranosyl)-5-methyl-4-pentafluorophenyl-pyrimidine-(2*H*)-one 193

The nucleoside (**191**) was co-evaporated with anhydrous pyridine (3 x 30 mL), dissolved in pyridine (90 mL) and cooled to 0 °C. To this was added 4-chlorophenyl phosphorodichloridate (8.90 mL, 54.68 mmol) with stirring. After 48 h, C_6F_5OH (15 g, 81.52 mmol) was added to the black reaction mixture. After 7 days the mixture was adsorbed onto silica. Flash chroamtography, eluting with EtOAc-hexane (1:1), afforded the title compound **193** (yield not recorded) as a cream-coloured solid; TLC [hexane-EtOAc (1:1)]: R_f 0.56 (lit.,[130] 0.55); ^{1}H NMR [($CDCl_3$]: δ 2.12 (s, 6 H, 2 x $CH_3C(O)$), 2.20 (s, 3 H, CH_3), 2.77 (m, 1 H, 2'-CH), 2.82 (m, 1 H, 2'-CH), 4.36 (m, 1 H, 4'-CH), 4.40 (m, 2 H, 5'-CH_2), 5.21 (m, 1 H, 3'-CH), 6.26 (t, 1 H *J* 6.3 Hz, 1'-CH) 7.86 ppm (s, 1H, 6-CH).

1-[2'-Deoxy-β-D-*erythro*-pentofuranosyl]-5-methyl-pyrimidine-2*H*-one 195

Nucleoside **193** was dissolved in a mixture of 1,4-dioxane-conc. NH_3 (aq) (3:1, 80 mL) and placed in a sealed bomb. The mixture was stirred during 4 h at r.t. After this time the product mixture was concentrated to give a viscous black oil which was dissolved in a saturated MeOH-NH_3 solution (60 mL) which was stirred in a sealed bomb at r.t. during 12 h. This mixture was then adsorbed onto silica. Flash chromatography, eluting firstly with EtOAc-MeOH (9:1) and then MeOH-EtOAc (1:1), afforded the title compound **195** (3.38 g, 51% over 4 steps); mp 217-220 °C (lit.,[173] 217-219 °C); ^{1}H NMR [($CDCl_3$]: δ 1.83 (s, 3 H, CH_3), 1.70-2.20 (m, 2 H, 2'-CH_2), 3.57 (m, 2 H, 5'-CH_2), 3.75 (m, 1 H, 4'-CH), 4.20 (m, 1 H, 3'-CH), 5.03 (br s, 1 H, 3' or 5'-OH), 5.22 (br s, 1 H, 3' or 5'-OH), 6.17 (t, 1 H, *J* 7 Hz, 1'-CH), 6.81 (br s, 1 H, NH), 7.32 (br s, 1 H, NH), 7.62 ppm (s, 1 H, 6-CH).

1-[2'-Deoxy-5'-*O*-(4,4'-dimethoxytrityl)-β-D-*erythro*-pentofuranosyl]-4-*N*-(*p*-*t*-butylphenoxyacetyl)-5-methyl-pyrimidine-2*H*-one 199

The nucleoside **195** (0.55 g, 2.30 mmol) was coevaporated with anhydrous pyridine (3 x 5 mL), dissolved in pyridine (5 mL) under Ar and cooled to 0 °C, was added TMSCl (1.2 mL, 9.46 mmol) and the mixture stirred during 1 h before again being cooled to 0 °C.

t-Butylphenoxyacetic anhydride (1.37 g, 3.44 mmol) was added as a solution in pyridine (2 mL). After 3 h 30 min, MeOH (0.5 mL) and H_2O (0.5 mL) were added to the reaction which was stirred for a further 20 min after which the solvent was evaporated. The residue was co-evaporated with toluene (2 x 5 mL), anhydrous pyridine (3 x 5 mL) and was suspended in pyridine (10 mL). DMTCl (0.86 g, 2.53 mmol) was added as a solution in anhydrous pyridine (2 mL). After 24 h, EtOH (1 mL) was added and the mixture was stirred for a further 30 min and thereafter concentrated to a red viscous oil. Flash chromatography, eluting firstly with hexane-EtOAc-Et_3N (40:60:1) and then EtOH-EtOAc-Et_3N (20:80:1) afforded the title compound **199** (0.42 g, 25% over two steps) as a yellow foam; mp 99-101 °C; TLC [EtOAc]: R_f 0.59; IR (KBr disc): ν_{max} 3400, 2960, 1706, 1660, 1606, 1564, 1514 cm^{-1}; ^1H NMR [$(CD_3)_2SO$]: δ 1.25 (s, 9 H, Bu^t), 1.61 (s, 3 H, 5-CH_3), 2.15-2.30 (m, 2 H, 2'-CH_2), 3.25 (m, 2 H, 5'-CH_2), 3.74 (s, 6 H, 2 x CH_3O), 3.96 (m, 1 H, 4'-CH), 4.32 (m, 1 H, 3'-CH), 5.00 (s, 2 H, $OCH_2C(O)N$), 5.39 (d, 1 H *J* 6.3 Hz, 3'-OH), 6.13 (t, 1 H *J* 6.3 Hz, 1'-CH), 6.80-7.35 (m, 17 H, 13 x CH (Ar) and 4 x CH (Ar, (BPA^t)), 7.89 (s, 1 H, 6-CH), 10.05 ppm (s, 1 H, NH); ^{13}C NMR [$(CD_3)_2SO$]: δ 20.5 (5-CH_3), 31.3 (Bu^t), 33.8 (Bu^tC), 40.8 (2'-CH_2), 55.0 (2 x CH_3O), 63.2 (5'-CH_2), 67.3 ($OCH_2C(O)N$), 70.0 (3'-CH), 85.9 (1'-CH, 4'-CH and Ar_3C), 113.3 (CH (Ar)), 114.0 (CH (Ar)), 125.3 (CH (Ar)), 126.0 (CH (Ar)), 126.8 (CH (Ar)), 127.7 (CH (Ar)), 128.0 (CH (Ar)), 128.2 (CH (Ar)), 128.9 (CH (Ar)), 129.7 (CH (Ar)), 135.3, 135.4 (2 x C), 143.2 (C-$C(CH_3)_3$), 144.6 (C), 153.0 (6-CH), 155.7 (4-C, $COCH_2C(O)N$ (tBPA)), 158.2 ppm (2-C, 2 x $COCH_3$, C(O)NH (tBPA)); Mass spectrum (FAB^+): *m/z* (I_r) 756 (M + Na, 26%), 303 (100%); Calcd for $C_{43}H_{47}N_3O_8$: (M + Na) 756.326. Found 756.328.

1-[2'-Deoxy-5'-*O*-(4,4'-dimethoxytrityl)-β-D-*erythro*-pentofuranosyl]-4-*N*-(*p*-*t*-butylphenoxyacetyl)-5-methyl-pyrimidine-2*H*-one-3'-*O*-succinate 200

Nucleoside **199** (0.09 g, 0.13 mmol) was dried by coevaporation with anhydrous pyridine (3 x 5 mL) and suspended in anhydrous pyridine (3 mL). DMAP (0.03 g, 0.25 mmol) was added followed by succinic anhydride (0.04 g, 0.40 mmol) as a solution in pyridine (1 mL), in 3 portions over 30 min. After 48 h, TLC indicated that the reaction was complete and the solvent was evaporated. The residue was suspended in CH_2Cl_2 (10 mL) and extracted with a 3% aqueous citric acid solution (2 x 15 mL) and dried over Na_2SO_4. The solvent was evaporated affording a brown gum which was used directly to derivatise LCAA-CPG. (0.05 g, 48% crude); TLC [EtOAc-MeOH-Et_3N (99:9:1)]: R_f 0.21; Mass spectrum (FAB^+): *m/z* (I_r) 856 (M + Na, 2%), 303 (100%).

1-Chloro-2-deoxy-3,5-di-*O*-*p*-toluoyl-α-D-*erythro*-pentofuranose 223

Acetyl chloride (1 mL) was added dropwise to MeOH (49 mL). This solution was then added in a single portion to a solution of 2'-deoxyribose (26.12 g, 150 mmol) dissolved in MeOH (450 mL) and the solution stirred for 10 min. The solution was then neutralised by addition of solid Ag_2CO_3 (9.80 g, 5.84 mmol) and the mixture filtered through Celite. The filtrate was washed with EtOAc and the combined filtrate was concentrated to a brown oil and dried under high vacuum.

The residual oil was dissolved in pyridine (160 mL) and the solution was cooled to 0 °C. *p*-Toluoyl chloride (64 mL, 48.73 mmol) was then added in four equal portions during 30 min resulting in the formation of a thick white precipitate. The reaction was stirred for a further 24 h, evaporated to dryness and the residue was dissolved in Et_2O (1500 mL). The organic phase was washed with of 0.1 M HCl (1 x 1500 mL), saturated $NaHCO_3$ (1 x 1500 mL), and the organic phase dried using Na_2SO_4 and evaporated giving a yellow oil.

The crude product from the above reaction was added to a mixture of glacial AcOH (300 mL) which had been saturated with HCl and the mixture was cooled to 0 °C. After stirring for 30 min, the precipitated product which had formed was collected by filtration and washed with Et_2O before drying under vacuum over P_2O_5 to yield the title compound **223** (39.50 g, 52% over 3 steps from 2'-deoxyribose); TLC [hexane-EtOAc (2:1)]: R_f = 0.40 (lit.,[130] 0.44); ^{1}H NMR [$(CD_3)_2SO$]: δ 2.40 (s, 3 H, CH_3), 2.41 (s, 3 H, CH_3), 2.70-2.91 (m, 2 H, 2'-CH_2), 4.63 (dd, 2 H, *J* 12.7 Hz, 5'-CH_2), 4.83-4.85 (m, 1 H, 4'-CH), 5.53-5.57 (m, 1 H, 3'-CH), 6.46 (d, 1 H, *J* 5.0 Hz, 1'-CH), 7.21-7.27 (m, 4 H, 4 x CH (Ar)), 7.87-8.00 ppm (m, 4 H, 4 x CH (Ar)).

6-Amino-2-chloro-9-(2'-deoxy-β-D-*erythro*-pentofuranosyl)-purine 214 and 6-Amino-2-chloro-7-(2'-deoxy-β-D-*erythro*-pentofuranosyl)-purine 225

A mixture of 2,6-dichloropurine (5.00 g, 26.45 mmol) and sodium hydride (1.13 g, 28.25 mmol, 60% dispersion in oil) in anhydrous MeCN (100 mL) was stirred at r.t. for 90 min under Ar. The chloro sugar **223** (10.28 g, 26.46 mmol) was added in four equal portions during 30 min. After 24 h, the reaction mixture was filtered through Celite. Evaporation of the solvent gave a mixture of **226** and **224** in a 2:1 ratio by ^{1}H NMR (7.82 g, crude). A saturated solution of MeOH-NH_3 (70 mL) was added to a portion of the mixture of dichloro nucleosides **226** and **224** (3.02 g, crude) and the mixture was heated in a bomb at 100 °C for 6 h. The product solution was evaporated concentrated and the residue purified by flash chromatography, eluting with CH_2Cl_2-MeOH (9:1) to give firstly compounds in the following order; N9 isomer (**214**) (0.40 g, 14% over 2 steps); mp > 300 °C (lit.,[161,162],

mp > 300 °C); TLC [MeOH-CH_2Cl_2 (1:4)]: R_f 0.38; ^{1}H NMR [$(CD_3)_2SO$]: δ 2.20-2.30 (m, 1 H, 2'-CH_2), 2.60-2.70 (m, 1 H, 2'-CH_2), 3.40-3.60 (m, 2 H, 5'-CH_2), 4.38 (m, 1 H, 3'-CH), 4.99 (t, 1 H, *J* 5.8 Hz, 5'-OH), 5.34 (d, 1 H, *J* 4.2 Hz, 3'-OH), 6.26 (t, 1H, *J* 6.4 Hz, 1'-CH), 7.86 (br s, 2 H, NH_2), 8.34 ppm (s, 1 H, 8-CH).

Further elution gave the N7 isomer (**225**) (0.10 g, 3% over 2 steps); mp > 300 °C (lit.,[161] mp > 300 °C); TLC [MeOH-CH_2Cl_2 (1:4)]: R_f 0.24; ^{1}H NMR [$(CD_3)_2SO$]: δ 2.28-2.33 (m, 2 H, 2'-CH_2), 3.42 (m, 2 H, 5'-CH_2), 3.90 (m, 1 H, 4'-CH), 4.39 (m, 1 H, 3'-CH), 5.19 (t, 1 H, *J* 4.6 Hz, 5'-OH), 5.41 (d, 1 H, *J* 4.5 Hz, 3'-OH), 6.30 (t, 1 H, 6.3 Hz, 1'-CH), 7.49 (br s, 2 H, NH_2), 8.56 ppm (s, 1 H, 8-CH); ^{13}C NMR [$(CD_3)_2SO$]: δ 41.2 (2'-C), 60.4 (5'-C), 69.3 (3'-C), 85.7 (1'-C), 88.0 (4'-C), 109.6 (5-C), 144.8 (8-C), 152.7 (6-C or 2-C), 153.2 (6-C or 2-C), 162.2 ppm (4-C).

6-Amino-2-benzyloxy-9-(2'-deoxy-β-D-*erythro*-pentofuranosyl)-purine 227

Sodium as a 50% dispersion in paraffin oil (0.19 g, 8.26 mmol) was added to anhydrous benzyl alcohol (15 mL) under Ar and the mixture was heated at 80 °C for 45 min to give a clear, colourless solution. The nucleoside **214** (0.45 g, 1.58 mmol) was then added and the solution was heated at 85 °C during 12 h. Silica (10 g) and MeCN (15 mL) were then added to the cooled, deep-red, reaction mixture and the solvent evaported. The resulting gel was added to a column of silica gel and the unreacted benzyl alcohol was eluted firstly using CH_2Cl_2 and then CH_2Cl_2-MeOH (95:5). Further elution with 10% MeOH-CH_2Cl_2 afforded the title compound **227** (0.40 g, 71%) as a yellow solid; mp 206-209 °C; TLC [MeOH-CH_2Cl_2 (1:4)]: R_f 0.52; IR (KBr disc): ν_{max} 3396, 3132, 2937, 1653, 1592 cm^{-1}; ^{1}H NMR [$(CD_3)_2SO$]: δ 2.22 (m, 1 H, 2'-CH_2), 2.71 (m, 1 H, 2'-CH_2), 3.50-3.58 (m, 2 H, 5'-CH_2), 3.85 (m, 1 H, 4'-CH), 4.41 (m, 1 H, 3'-CH), 5.07 (t, 1 H, *J* 5.6 Hz, 5'-OH), 5.31-5.33 (m, 3 H, CH_2 and 3'-OH), 6.26 (t, 1 H, *J* 6.4 Hz, 1'-CH), 7.31-7.47 (m, 7 H, 5 x CH (Ar), NH_2), 8.16 ppm (s, 1 H, 8-CH); ^{13}C NMR [$(CD_3)_2SO$]: δ 39.2 (2'-CH_2), 62.1 (5'-CH_2), 67.9 (CH_2), 71.2 (3'-CH), 83.7 (1'-CH), 88.0 (4'-CH), 116.0 (5-C), 127.9 (CH (Ar)), 128.0 (CH (Ar)), 128.5 (CH (Ar)), 137.6 (C (Bn)), 138.6 (8-CH), 150.8 (4-C), 157.0 (6-C) and 161.2 ppm (2-C). Mass spectrum (ES$^+$): *m/z* (I_r) 358 (M + H, 100%), 113 (35%); Calcd for $C_{17}H_{19}N_5O_4$: 410.008. Found 410.009; Calcd for $C_{17}H_{19}N_5O_4$: C, 57.1%; H, 5.3%; N, 19.6%. Found 57.2%; H, 5.3%; N, 19.5%.

2'-Deoxyisoguanosine 217

To nulceoside **227** (0.27 g, 0.75 mmol) dissolved in MeOH (70 mL) was added 5% Pd on C (0.22 g) and the mixture was hydrogenolysed at r.t. for 6.5 h using a hydrogen baloon. The product mixture was filtered through Celite after which activated charcoal (1 g) was

added and the mixture was boiled for 10 min. The charcoal was removed by filtration before concentration of the product mixture affording the title compound **217** (0.11 g, 53% yield) as an off-white solid; decomposed > 230 °C (lit.,[158] decomposed > 230 °C); ^{1}H NMR [$(CD_3)_2SO$]: δ 2.14 (m, 1 H, 2'-CH_2), 2.49 (m, 1 H, 2'-CH_2), 3.53 (m, 2 H, 5'-CH_2), 3.83 (m, 1 H, 4'-CH), 4.33 (m, 1 H, 3'-CH), 5.27 (br s, 1 H, 3'-OH), 5.60 (br s, 1 H, 5'-OH), 6.09 (t, 1 H, *J* 7.9 Hz, 1'-CH), 7.70 (br s, 2 H, NH_2), 7.94 (s, 1 H, 8-CH), 10.70 ppm (br s, 1 H, NH); ^{13}C NMR [$(CD_3)_2SO$]: δ 39.3 (2'-C), 62.0 (5'-C), 71.1 (3'-C), 83.8 (1'-C), 88.0 (4'-C), 109.6 (5-C), 137.7 (8-C), 152.5 (C6, tentative), 156.1 ppm (2-C, tentative). Lit.[168].

6-Amino-2-allyloxy-9-(2'-deoxy-β-D-*erythro*-pentofuranosyl)-purine 233

Sodium (0.64 g, 27.83 mmol) as a 50% as 50% dispersion in paraffin oil was added to allyl alcohol (14 mL) in three portions over 5 min at 0 °C under Ar. After 1 h, the nucleoside **214** (0.76 g, 2.65 mmol) was added and the mixture was heated at 88 °C during 12 h after which time the mixture was concentrated to give a red oil which was purified by flash chromatofraphy using a gradient elution using CH_2Cl_2 containing MeOH (0-20%) affording the title compound **222** (0.50 g, 62%) as a cream coloured solid.; mp 170-173 °C; TLC [MeOH-CH_2Cl_2 (1:4)]: R_f 0.47; IR (KBr disc): ν_{max} 3371, 3336, 2942, 2874, 1662, 1610 cm^{-1}; ^{1}H NMR [$(CD_3)_2SO$]: δ 2.22 (m, 1 H, 2'-CH_2), 2.71 (m, 1 H, 2'-CH_2), 3.53 (m, 2 H, 5'-CH_2), 3.84 (m, 1 H, 4'-CH), 4.38 (m, 1 H, 3'-CH), 4.74 (m, 2 H, CH_2O), 5.04 (br s, 1 H, 5'-OH), 5.17-5.38 (m, 3 H, CH_2, 3'-OH), 6.03 (m, 1 H, CH), 6.22 (t, 1 H *J* 6.9 Hz, 1'-CH), 7.32 (br s, 2 H, NH_2), 8.13 ppm (s, 1 H, 8-CH); ^{13}C NMR [$(CD_3)_2SO$]: δ 40.0 (2'-C), 62.0 (5'-C), 66.8 (CH_2O), 71.0 (3'-C), 83.5 (1'-C), 87.8 (4'-C), 115.7 (5-C), 116.9 (CH_2), 134.0 (CH), 138.4 (8-C), 150.7 (4-C), 156.9 (6-C), 160.9 ppm (2-C). Mass spectrum (ES$^+$): *m/z* (I_r) 308 (M + H, 100%); Calcd for $C_{13}H_{17}N_5O_4$: (M + H) 308.136. Found 308.135; Calcd for $C_{13}H_{17}N_5O_4$: C, 51.0%; H, 5.2%, N, 22.9%. Found 50.7%; H, 5.5%; N, 22.5%.

6-Amino-2-benzyloxy-6-*N*-[1-(dimethylamino)-ethylidene]-9-(2'-deoxy-5'-*O*-(4,4'-dimethoxytrityl)-β-D-*erythro*-pentofuranosyl)-purine 235

Nucleoside **227** (0.55 g, 1.54 mmmol) was coevaporated with anhydrous pyridine (3 x 10 mL), dissolved in anhydrous MeOH (3 mL) and placed under Ar. To this was added *N,N*-dimethylacetamide dimethyl acetal (0.68 mL, 4.65 mmol) and the mixture was heated at 40 °C. After 26 h, H_2O (0.3 mL) was added and the mixture stirred for a further 10 min. The mixture was concentrated, the residue co-evaporated with anhydrous pyridine (3 x 3 mL) and dissolved in pyridine (15 mL) under Ar. DMTCl (0.64 g, 1.89 mmol) was added and the solution stirred during 12 h after which time MeOH (0.5 mL) was added and the mixture stirred for a further 10 min. The mixture was concentrated and purified by flash

chromatography eluting firstly with EtOAc-Et_3N (100:1) and then MeOH-EtOAc-Et_3N (5:95:1, then 10:90:1) affording a yellow foam. This was precipitated from CH_2Cl_2 (3 mL) into cold hexane-Et_2O (2:1, 700 mL). The product was dissolved in CH_2Cl_2 (10 mL) and dried under high vacuum at 40 °C to give the title compound **235** (0.99 g, 88% yield over 2 steps) as a white solid; mp 102-105 °C; TLC [MeOH-EtOAc-Et_3N (10:90:1)]: R_f 0.37; IR (KBr disc): ν_{max} 3413, 2929, 1606, 1564 cm^{-1}; ^1H NMR [$(CD_3)_2SO$]: δ 2.01 (s, 3 H, CH_3C), 2.27 (m, 1 H, 2'-CH_2), 2.81 (m, 1 H, 2'-CH_2), 3.08 (s, 8 H, 5'-CH_2, $(CH_3)_2N$), 3.68 (s, 6 H, 2 x CH_3O), 3.96 (m, 1 H, 4'-CH), 4.45 (m, 1 H, 3'-CH), 5.17 (m, 2 H, CH_2), 5.34 (m, 1 H, 3'-OH), 6.32 (m, 1 H, 1'-CH), 6.68-7.35 (m, 18 H, 13 x CH (Ar), 5 x CH (Ar)), 8.13 ppm (s, 1 H, 8-CH); ^{13}C NMR [$(CD_3)_2SO$]: δ 17.2 (CH_3C), 37.1, 38.2, ($(CH_3)_2N$), 38.4 (2'-CH_2), 54.96, 54.99 (2 x CH_3O), 64.5 (5'-CH_2), 68.1 (CH_2), 70.8 (3'-CH_2), 83.5 (1'-CH), 85.4 (Ar_3C), 86.0 (4'-CH), 113.1 (CH), 122.3 (5-C), 126.6 (CH), 127.7 (CH (Ar)), 127.8 (CH (Ar)), 128.3 (CH (Ar)), 129.65 (CH (Ar)), 129.72 (CH (Ar)), 135.5 (C), 135.7 (C), 137.3 (C (Ar)), 140.2 (8-CH), 145.0 (C (Ph)), 152.0 (4-C), 157.97 ($COCH_3$), 158.01 ($COCH_3$), 160.7, 160.9, 161.3 ppm (C-2, C-6, C=N); Mass spectrum (FAB^+): *m/z* (I_r) 729 (M + H, 50%), 303 (100%); Calcd for $C_{42}H_{44}N_6O_6$: (M + H) 729.340. Found 729.343; Calcd for $C_{42}H_{44}N_6O_6$: C, 69.2%; H, 6.0%; N, 11.5%. Found C, 68.7%; H, 6.0%; N, 11.5%.

6-Amino-2-allyloxy-6-*N*-[1-(dimethylamino)-ethylidene]-9-(2'-deoxy-5'-*O*-(4,4'-dimethoxytrityl)-β-D-*erythro*-pentofuranosyl)-purine 238

The nucleoside **233** (0.89 g, 2.90 mmol) was co-evaporated with anhydrous pyridine (3 x 10 mL) and dissolved in anhydrous MeOH (2 mL) under Ar. To this, *N,N*-dimethylacetamide dimethylacetal (1.7 mL, 11.63 mmol) was added and the mixture stirred at r.t. for 47 h after which time water (0.3 mL) was added and the mixture stirred for a further 10 min. The product mixture was concentrated, the residue co-evaporated with anhydrous pyridine (3 x 3 mL) and dissolved in pyridine (40 mL) under Ar. DMTCl (1.08 g, 3.19 mmol) was added in two portions over 20 min and the solution stirred for 110 min at r.t. after which MeOH (0.05 mL) was added and the mixture stirred for a further 10 min. The product mixture was concentrated and the residue purified by flash chromatography, eluting with MeOH-EtOAc-Et_3N (5:95:1 and then 10:90:1) affording a yellow foam which was precipitated from CH_2Cl_2 (3 mL) into hexane-Et_2O (2:1, 750 mL). The product was dissolved in CH_2Cl_2 (3 mL) and dried under high vacuum to give the title compound **238** (1.33 g, 68% yield over 2 steps) as a white solid; mp 98-101 °C; TLC [MeOH-EtOAc-Et_3N (10:90:1)]: R_f 0.26; IR (KBr disc): ν_{max} 3265, 2929, 1564, 1505 cm^{-1}; ^1H NMR [$(CD_3)_2SO$]: δ 2.04 (s, 3 H, CCH_3), 2.32 (m, 1 H, 2'-CH_2), 2.87 (m, 1 H, 2'-CH_2), 3.00-3.20 (m, 8 H, $N(CH_3)_2$ and 5'-CH_2), 3.69 (s, 3 H, CH_3O), 3.70 (s, 3 H, CH_3O), 3.98 (m, 1

H, 4'-CH), 4.49 (m, 1 H, 3'-CH), 4.69 (m, 2 H, CH_2O), 5.16-5.41 (m, 3 H, CH_2 and 3'-OH), 6.00 (m, 1 H, CH), 6.73-7.34 (m, 13 H, 13 x CH (Ar)), 8.14 ppm (s, 1 H, 8-CH); ^{13}C NMR [$(CD_3)_2SO$]: δ 17.4 (CH_3C), 38.2, 38.4 $(CH_3)_2N$), 38.6 (2'-CH_2), 55.1 (CH_3O), 55.2 (CH_3O), 64.6 (5'-CH_2), 67.2 (CH_2), 71.0 (3'-CH), 83.5 (1'-CH), 85.6 (Ar_3C), 86.0 (4'-CH), 113.2 (CH (Ar)), 113.2 (CH (Ar)), 117.1 (CH_2), 122.4 (5-C), 126.8 (CH (Ar)), 127.9 (CH (Ar)), 129.8 (CH (Ar)), 129.9 (CH (Ar)), 134.1 (CH), 135.7 (C), 135.8 (C), 140.2 (8-CH), 145.1 (C (Ph)), 152.2 (C-4), 158.2 (2 x $COCH_3$), 160.8, 161.1, 161.4 ppm (C2, C6, C=N); Mass spectrum (FAB$^+$): *m/z* (I_r) 679 (M + H, 55%), 303 (100%); Calcd for $C_{38}H_{42}N_6O_6$: (M + H) 679.324. Found 679.326; Calcd for $C_{38}H_{42}N_6O_6$: C, 67.3%; H, 6.2%; N, 12.4%. Found C, 66.8%; H, 6.0%; N, 12.3%.

6-Amino-2-benzyloxy-6-*N*-[1-(dimethylamino)-ethylidene]-9-(2'-deoxy-5'-*O*-(4,4'-dimethoxytrityl)-β-D-*erythro*-pentofuranosyl)-purine-3'-*O*-(2-cyanoethyl)-*N*,*N*-diisopropylphosphoramidite 236

Nucleoside **235** (0.20 g, 0.28 mmol) was dried by coevaporation with anhydrous pyridine (3 x 5 mL) and suspended in anhydrous THF (5 mL) under Ar. DIPEA (0.19 mL, 1.09 mmol) followed by 2-cyanoethyl-*N*,*N*-diisopropyl chlorophosphoramidite (0.09 mL, 0.40 mmol) were then added with stirring. After 40 min a white precipitate had formed and the product mixture was concentrated and the residue purified by flash chromatography eluting with EtOAc-CH_2Cl_2-Et_3N (50:50:1) affording an oil which was precipitated five times from CH_2Cl_2 (3 mL) into cold hexane (300 mL). The product was dissolved in CH_2Cl_2 (3 mL) and dried under high vacuum at 30 °C giving the title compound **236** (0.09 g, 33%) as a colourless viscous oil; mp 66-69 °C; TLC [EtOAc-CH_2Cl_2-Et_3N (50:50:1)]: R_f 0.28; IR (KBr disc): ν_{max} 3551, 3473, 2962, 2961, 1606, 1570 cm^{-1}; ^1H NMR [$(CD_3)_2SO$]: δ 0.95-1.23 (m, 12 H, 4 x CH_3), 2.03 (s, 3 H, CH_3C), 2.62 (t, 1 H *J* 5.8 Hz, CH_2CN), 2.74 (t, 1 H *J* 5.8 Hz, CH_2CN), 2.95-3.25 (m, 8 H, $(CH_3)_2N$, 2'-CH_2), 3.40-3.80 (m, 12 H, 2 x CH_3O, $POCH_2$, 2 x NCH, 5'-CH_2), 4.10 (m, 1 H, 4'-CH), 4.82 (m, 1 H, 3'-CH), 5.11-5.24 (m, 2 H, CH_2), 6.31 (m, 1 H, 1'-CH), 6.68-7.35 (m, 18 H, 13 x CH (Ar), 5 x CH (Ph)), 8.16, 8.17 ppm (2 s, 1 H, 8-CH); ^{31}P NMR [$(CD_3)_2SO$]: δ 147.8, 148.5 ppm; ^{13}C NMR [$(CD_3)_2SO$]: δ 17.2 (CH_3C), 19.8 (CH_2CN), 24.1, 24.2 (4 x CH_3), 37.0, 37.1 (2'-CH_2), 37.6, 38.1 ($(CH_3)_2N$), 42.5, 42.7 (2 x NCH), 54.9 (2 x CH_3O (Ar)), 58.2, 58.3, 58.4, 58.6 ($POCH_2CH_2CN$), 63.8 (5'-CH_2), 68.1, 68.2 (CH_2 (Ph)), 73.3, 73.6 (3'-CH), 83.6 (1'-CH), 84.6, 84.8 (4'-CH), 85.4 (Ar_3C), 113.0 (CH (Ar)) 118.9, 119.0 (CN), 122.3 (5-C), 126.6, 127.6, 127.7, 128.3, 129.6 (CH (Ar)), 135.4 (2 x C (Ar)), 137.15, 137.22 (C (Ph)), 140.5 (8-CH), 144.8 (C (Ph)), 151.8 (4-C), 158.0 (2 x $COCH_3$), 160.7, 160.9, 161.4 ppm (2-C, 6-C, C=N); Mass spectrum (APCI$^+$): *m/z* (I_r) 929 (M + H, 38%), 303 (100%); Calcd for

$C_{51}H_{61}N_8O_7P$: (M + Na) 951.430. Found 951.429. Calcd for $C_{51}H_{61}N_8O_7P$: C, 66.0%; H, 6.6%; N, 12.1%; P, 3.3. Found C, 65.7%; H, 6.5%; N, 12.0; P, 3.2%.

6-Amino-2-allyloxy-6-*N*-[1-(dimethylamino)-ethylidene]-9-(2'-deoxy-5'-*O*-(4,4'-dimethoxytrityl)-β-D-*erythro*-pentofuranosyl)-purine-3'-*O*-(2-cyanoethyl)-*N*,*N*-diisopropylphosphoramidite 239

The nucleoside **238** (0.22 g, 0.32 mmol) was coevaporated with anhydrous pyridine (3 x 5 mL) and suspended in anhydrous THF (4 mL) under Ar. To this solution was added DIPEA (0.22 mL, 1.26 mmol) followed by 2-cyanoethyl-*N*,*N*-diisopropyl chlorophosphoramidite (0.11 mL, 0.49 mmol). After 1 h a white precipitate had formed and the product mixture was concentrated and the residue purified by flash chromatography eluting with EtOAc-CH_2Cl_2-Et_3N (50:50:1) affording an oil which was precipitated from CH_2Cl_2 (3 mL) into cold hexane (300 mL). The product was dissolved in CH_2Cl_2 (3 mL) and dried under high vacuum at 30 °C giving the phosphoramidite **239** (0.18 g, 64%) as a colourless viscous oil; mp 59-62 °C; TLC [CH_2Cl_2-EtOAc-Et_3N (50:50:1)]: R_f 0.13; IR (KBr disc): ν_{max} 3411, 3056, 2968, 1606, 1568 cm^{-1}; ^1H NMR [$(CD_3)_2SO$]: δ 0.96-1.19 (m, 12 H, 4 x CH_3), 2.03 (s, 3 H, CH_3), 2.63 (t, 1 H *J* 5.9 Hz, CH_2CN), 2.75 (t, 1 H *J* 5.8 Hz, CH_2CN), 2.95-3.25 (m, 8 H, $(CH_3)_2N$, 2'-CH_2), 3.40-3.80 (m, 10 H, 2 x CH_3O, $POCH_2$, 2 x NCH, 5'-CH_2), 4.1 (m, 1 H, 4'-CH), 4.62-4.80 (m, 3 H, CH_2O, 3'-CH), 5.16-5.32 (m, 2 H, CH_2), 5.75 (m, 1 H, CH), 6.31 (m, 1 H, 1'-CH), 6.70-7.30 (m, 13 H, 13 x CH (Ar)), 8.13, 8.14 ppm (2 s, 1 H, 8-CH); ^{31}P NMR [$(CD_3)_2SO$]: δ 147.9, 148.6 ppm; ^{13}C NMR [$(CD_3)_2SO$]: δ 17.1 (CH_3C), 19.4 (CH_2CN), 19.76 (CH_2CN), 19.8 (CH_2CN), 19.9 (CH_2CN), 24.1, 24.18, 24.23, 24.27, 24.29, 24.34, 24.39 (4 x CH_3), 36.9, 37.0 (2'-CH_2), 37.4, 38.1 ($(CH_3)_2N$), 42.5, 42.7 (2 x NCH), 54.92, 54.95, 54.96 (2 x CH_3O), 58.1, 58.3, 58.4, 58.5 ($POCH_2CH_2CN$), 63.6 (5'-CH_2), 67.1 (CH_2O), 72.9, 73.1, 73.3, 73.6, (3'-CH), 83.5 (1'-CH), 84.4, 84.5, 84.7 (4'-CH), 85.5 (Ar_3C), 113.0 (CH (Ar)), 116.9, 117.0 (CH_2), 118.7, 119.0 (CN), 122.2, 122.3 (5-C), 126.6, 127.6, 127.63, 129.5, 129.6, 129.64 (CH (Ar)), 133.3, 133.8, 133.81 (CH), 135.37, 135.43, 135.45 (2 x C (Ar)), 140.2, 140.4 (8-CH), 144.8 (PhC), 151.7, 151.8 (4-C), 157.96, 157.97, 158.0 (2 x $COCH_3$), 160.5, 160.9, 161.3 ppm (2-C, 6-C, C=N); Mass spectrum (APCI$^+$): *m/z* (I_r) 879 (M + H, 50%), 303 (100%); Calcd for $C_{47}H_{59}N_8O_7P$: (M + Na) 901.414. Found 901.410.

5.2 Solid phase synthesis

5.2.1 General methods

Automated solid phase oligonucleotide synthesis was carried out on a 0.2 or 1 μmol scale using a Beckmann Oligo 1000 DNA Synthesiser following the standard protocol recommended by the manufacturer and with minor modifications, where stated. 2'-

Deoxyribonucleotide cyanoethylphosphoramidites containing T, A^{Bz}, $C^{Bz \text{ or } Ac}$ and G^{Bz} bases and conventional, derivatised LCAA-CPG supports were purchased from Cruachem and from Beckmann. The C^{BPA^t} phosphoramidite and BPA^t anhydride CAP A solution were purchased from Perseptive Biosystems. The nucleoside 3'-*O*-succinates (dG^{Bui}, dA^{Bz}, dC^{Bz}, dT) and LCAA-CPG were purchased from the Sigma Chemical Company. In all instances where C^{BPA^t} phosphoramidite was used, BPA^t anhydride CAP A solution was also used as recommended by the manufacturer. Molecular sieves (3 angstrom) were purchased from Beckmann and were used to dry all phosphoramidites synthesised in this thesis work. NAP-10 columns were purchased from Pharmacia. All sequences were synthesised in "DMT-OFF" mode, unless otherwise stated.

5.2.2 Standard cleavage and deprotection of oligonucleotides

Where conventional, succinyl-linked CPG (also referred to as normal CPG) was used for DNA synthesis, conc. NH_3 (aq) at 55 °C during 16 h was used to cleave the sequence from the support and for removal of all protecting groups. Where SLCPG was used for DNA synthesis the following cleavage/deprotection condition sets were used, either in whole or in part, depending on sequence.

5.2.3 Methods of cleaving oligonucleotides from SLCPG support

(a) TBAF method: Protected oligonucleotides were removed from SLCPG after automated synthesis by treating the solid support with 0.5 M TBAF-0.5 M AcOH solution in THF (0.4 mL) during 1 min.[92]

(b) TFA method: The oligonucleotide was cleaved from the support by treating it with a 2% TFA/CH_2Cl_2 solution (1 mL) with shaking for 1 h in a screw-cap vial after which the solvent was evaporated under reduced pressure. The crude sequence was then suspended in H_2O (standard sequences) or a 50% MeCN-H_2O mixture (for oligonucleotide-drug conjugates) from which the CPG was filtered through a Whatman filter (0.45 μm). The solvent was evaporated using a Speed Vac under reduced pressure affording the crude oligonucleotide for HPLC analysis/purification.

(c) $Et_3N.3HF$ method: The oligonucleotide was cleaved from the support using $Et_3N.3HF$ during 15 min in a screw-cap sample tube with shaking. This mixture was then filtered through a Whatman filter (0.45 μm), washing with a H_2O-MeCN (1:1) mixture. Solvent was evaporated in a Speed Vac affording the crude oligonucleotide for HPLC analysis/purification.

5.2.4 Methods of removing cyanoethyl protecting groups

(a) TBAF method: Following support cleavage (Section 5.2.3 (a)) oligonucleotides were deprotected by heating the TBAF-AcOH solution from Section 5.2.3 (a) above at 50 °C during 2 h. Oligonucleotide solutions thus formed were desalted using a SepPak (Millipore) column eluting firstly with H_2O. Oligonucleotides were eluted subsequently using H_2O-CH_3CN (1:1). Fractions containing the products were lyophilised to give the crude oligonucleotides.

(b) Bu^tNH_2 method: The SLCPG-bound sequence was placed in a screw-cap vial to which was added Bu^tNH_2 (1 mL) and the tube was shaken overnight (15.5 h). The support was washed with MeCN (3 x 1 mL), Et_2O (2 x 1 mL) and dried under vacuum.

5.2.5 Removal of BPA^t protecting groups using ethanolamine

The support-bound sequence was then treated with ethanolamine (1 mL) with vigorous shaking during 30 min. The SLCPG was then washed with MeCN (2 x 1 mL), MeOH (2 x 1 mL) and Et_2O (2 x 1 mL) and dried under vacuum.

5.2.6 Trityl assay

Step-wise and individual coupling efficiencies and synthesis yields were determined spectrophotometrically by DMT cation assay according to a literature procedure.[174]

5.2.7 Support load determination

Determined according to a literature procedure.[49]

5.2.8 HPLC analysis and purification of oligonucleotides

HPLC analysis was carried out using a Hewlett Packard Series 1100 instrument with the UV detection wavelength set at 260 nm. Gradient elution was performed on a HPLC Technology C18 Reversed Phase Column (250 x 4.6 mm) at a flow rate of 1 mL min^{-1} where solvent system A was mixed with solvent system B. Solvent system A was composed of 1 M aqueous triethylammonium acetate (TEAA, 10%) and CH_3CN (2%) at pH 7.0, and solvent system B was composed of 1 M aqueous TEAA (10%) and CH_3CN (80%) at pH 7.0. Semi-preparative purification was carried out by dissolving the impure oligonucleotide in H_2O (2 mL) and injecting aliquots (typically 0.2 to 0.3 mL) using a 1 mL injection loop and syringe. All solvent and column conditions were as above, except that the UV detection wavelength was 285 nm. Fractions were collected manually in Eppendorf tubes. Appropriate fractions were combined and an analytical run was made (as

outlined above) to assess oligonucleotide purity. The volatile TEAA salt was removed by co-evaporation as follows:

1. Evaporation of the mobile phase was performed at reduced pressure in a Speedvac.

2. The residue from (1.) was suspended in H_2O (1 mL) and evaporated under reduced pressure.

3. The residue from (2.) was suspended in 50% ethanol solution (1 mL) and evaporated under reduced pressure.

5.2.9 Electrospray mass spectrometric analysis of oligonucleotides

Desalted samples were analysed by ESMS in negative ion mode. Occasionally, samples were desalted further by passing through a NAP-10 column according to the manufacturer's instructions[175] prior to ESMS analysis. Samples were dissolved in weakly basic NH_3 (aq) solutions for ESMS analysis. However, the following solutions were used for analysis of the oligonucleotide-drug conjugates:

a. 5% NH_3-AcOH

b. 5% TEAA

c. 40% Bu^tOH-H_2O

5.3 Synthesis of oligonucleotides and conjugates

5.3.1 Synthesis of DMTdT derivatised SLCPG support (95)[93]

A solution of freshly sublimed imidazole (70 mg, 1.03 mmol) and *bis*-(trifluoromethanesulfonyl)diisopropylsilane (0.15 mL, 0.5 mmol) in a 1:1 mixture of anhydrous MeCN-DMF (3 mL) was prepared under inert argon atmosphere at -40 °C using an MeCN dry-ice bath and was stirred during 10 min. A solution of DMTdT (**97**) (253 mg, 0.47 mmol) and imidazole (34 mg, 0.5 mmol) in a 1:1 mixture of anhydrous MeCN-DMF (3 mL) was also prepared under inert Ar atmosphere and was added to the first solution in three portions during 15 min. The solution was allowed to warm to r.t. and was then added to a mixture of support (**99**) (450 mg, 55 μmol g^{-1}), 2,6-lutidine (0.2 mL, 1.72 mmol) and DMAP (35 mg, 0.29 mmol) under inert Ar atmosphere which was shaken gently during 19 h. The CPG support was collected by filtration and washed with DMF, MeOH, Et_2O affording the desired DMTdT derivatised SLCPG support (**95**) (load = 30 μmol g^{-1}).

5.3.2 SLCPG cleavage using $Et_3N.3HF$

The sequence d(T_8) was synthesised on a 1 μmol scale on SLCPG. Four portions of support (~ 2 mg) were each treated with $Et_3N.3HF$ (0.1 mL), with shaking, for the time periods shown in Table 5.1. The samples were then diluted with MeCN-H_2O (to 1 mL) and the support was filtered. Samples were evaporated to dryness and suspended in H_2O (1500 μL). Samples were further diluted by a third and analysed by UV (260 nm) (Table 5.1).

Table 5.1 Cleavage of SLCPG support (**95**) using $Et_3N.3HF$.

Time	Mass SLCPG (mg)	$Abs^{260\ nm}$	Abs per mass of SLCPG	Support Cleavage (%)
5 min	2.36	1.203	0.5097	93
15 min	2.66	1.333	0.5011	94
30 min	2.23	1.170	0.5247	95
16.5 h	2.43	1.338	0.5506	100

Sample calculation of % support-cleavage: 0.5097/0.5506 x 100% = 93%

5.3.3 Stability of CPG towards treatment with $Et_3N.3HF$ during 10 min

$Et_3N.3HF$ (1 mL) was added to CPG (57.2 mg) with shaking for 10 min. The CPG was then washed with H_2O-MeCN (1:1, 2 mL), MeOH (1 x 2 mL), Et_2O (1 x 2 mL). The remaining CPG was dried under vacuum and weighed (37.5 mg, 66% recovery).

5.3.4 Efficiency of 2%TFA for SLCPG cleavage during 1 h

Sequence **101** d(T_6) was synthesised on SLCPG support. The fully protected sequence was cleaved from the support using the standard TBAF-AcOH procedure (Section 5.2.3 (a)) as well as 2% TFA/CH_2Cl_2 during 1 h (Section 5.2.3 (b)). After evaporating the crude oligonucleotide solutions to dryness, each was resuspended in 400 μL H_2O, of which a 100 μl aliquot was diluted to 1500 μl with H_2O. The results of UV analysis (260 nm) are shown in Table 5.2. The UV absorbance per weight of SLCPG support was determined, which was assumed to be directly proportional to the concentration of oligonucleotide in each solution.

Table 5.2 Cleavage of SLCPG support (**95**) using 2% TFA/CH_2Cl_2 during 1 h.

Conditions	Mass SLCPG (mg)	$Abs^{260\ nm}$	Abs/mass SLCPG
2%TFA/CH_2Cl_2	7.9	0.466	0.059
1:1 TBAF/AcOH	7.3	0.439	0.060

The percentage efficiency of the 2%TFA/CH_2Cl_2 method relative to the standard TBAF-AcOH method was calculated to be 98% as follows:

$$0.059/0.060 \times 100\% = 98\%$$

5.3.5 Stability of SLCPG support (95) towards ethanolamine

Portions of DMTdT derivatised SLCPG support (**95**) (28 μmol g^{-1}) were treated with ethanolamine for the time periods shown in Table 5.3 and were then washed with MeCN (2 x 1 mL), MeOH (2 x 1 mL) and Et_2O (2 x 1 mL). The trityl load of each portion was then assessed (Section 5.2.7) and the loss of support due to treatment with ethanolamine was calculated by comparison with the load of the untreated sample.

Table 5.3 Treatment of SLCPG support (**95**) with ethanolamine.

Time	Mass SLCPG (mg)	Vol $HClO_4$ (mL)	Abs[a]	Support load (μmol g^{-1})	Support Cleavage (%)
0	-	-	-	28.0	0
20 min	3.05	6.0	0.940	26.4	6
100 min	2.93	5.8	0.845	23.9	15
3 h	1.38	2.8	0.624	18.1	35
24 h	0.51	1.5	0.115	4.84	83

Typical support-load calculation for 20 min treatment: (6.0 x 0.94 x 14.3)/3.05 = 26.4 μmol g^{-1}.

[a] Absorbace values are the average of two measurements.

5.3.6 Stability of SLCPG support (95) towards 3%$CHCl_2CO_2H/CH_2Cl_2$

A portion of DMTdT-derivatised SLCPG support (**95**) (28 μmol g^{-1}) was treated with 3% $CHCl_2CO_2H$-CH_2Cl_2 during 10 min. The support was then washed with MeOH (3 x 2 mL) and Et_2O (3 x 2 mL) and dried under vacuum. A portion (13.4 mg) was then treated to standard TBAF/AcOH cleavage conditions (Section 5.2.3. (a)) during 1 min. To this was added MeOH (2000 μl) and the mixture was shaken and sonicated. A UV spectrum (200-500 nm) of a blank, sample matrix solution (2000 μl MeOH, 415 μl of 0.5 M AcOH, 0.5 M TBAF in THF) was run as background. The supernatant liquid of the sample solution was then run over the same range (λ_{max} 267 nm, A 0.627). An aliquot (5 μl) of a stock solution of 2'-deoxythymidine (41.24 mM) was added to a 2 mL volumetric flask which made to the mark with the sample solution which was then analysed by UV (λ_{max} 267 nm, ε 8759.5 $Lmol^{-1}cm$, lit.,[174] 8800 $Lmol^{-1}cm$). The load of dT nucleoside on SLCPG (28 μmol g^{-1})

after treatment with 3% $CHCl_2CO_2H/CH_2Cl_2$ solution was found to be 15.6 μmol g^{-1} (56% yield: 44% SLCPG support cleavage (**95**) through treatment with 3% $CHCl_2CO_2H/CH_2Cl_2$ solution).

5.3.7 Synthesis of sequence 108

DNA synthesis of oligonucleotide **108** was typically carried out on a 1 μmol scale using phosphoramidite **103** (0.1 M) and succinyl CPG support. The support bound sequence was cleaved and deprotected using conc. NH_3 (aq) solution (Section 5.2.2). Two peaks were purified by HPLC and their identities confirmed as **108** (t_R 14.8 min; MW found 2535.3, calcd 2535.5) and **109** (t_R 15.5 min; MW found 2578.2, calcd 2578.5). An overall yield of 76% was calculated for the entire oligonucleotide synthesis, based on trityl yield. A stepwise yield of 98% was observed. Phosphoramidite **103** was coupled in 83% yield.

Similarly, **108** were synthesised using SLCPG support. Reagents including ethanolamine, DIPA and 20% Et_3N-pyridine were used under varied conditions in attempts to remove the cyanoethyl groups from the support-bound sequences to give **116** (Table 5.4). In all cases, standard TBAF-AcOH cleavage of the support-bound sequence **116** was used to give **108**.

Table 5.4 Reagents used for cyanoethyl group removal from the SLCPG support-bound sequence **116**.

Conditions	Deprotection
Ethanolamine 10, 15, 20, 40 min, r.t.	Incomplete
DIPA (dry) 14 h, 24 h, r.t.	Incomplete
DIPA (dry) 17 h, 60 °C	Complete
20% Et_3N-pyridine 2 h, r.t.	Incomplete
20% Et_3N-pyridine 24 h, r.t.	Complete

Where deprotection was incomplete, peaks with t_R ~ 20 min were present in the HPLC profile. Treatment of these samples with conc. NH_3 (aq) afforded sequences **108** and **109**, in a ratio of 54:46 by peak area, the identities of which were confirmed by coinjection with authentic samples.

Optimal conditions for the synthesis of sequence **108** involved cleavage of the cyanoethyl groups from the support-bound sequence to give **116** using Bu^tNH_2 at r.t. during 15.5 h (Section 5.2.4 (b)) followed by detatchment of the sequence (**108**) from the support using 2% TFA/CH_2Cl_2 (Section 5.2.3 (b)). These conditions were used for subsequent conjugate synthesis.

5.3.8 Stability of SLCPG support (95) towards Bu^tNH_2

A portion of DMTdT-derivatised SLCPG support (**95**) (~10 mg) was treated, with shaking, to anhydrous Bu^tNH_2 (0.5 mL) during 15.5 h. After this time the SLCPG support was washed with MeCN (2 x 1 mL), MeOH (2 x 1 mL) and Et_2O (2 x 1 mL). The trityl load of two portions of this support (2.24 mg, 2.34 mg) was assessed (Section 5.2.7) and the average load was found to be 53.7 μmol g^{-1} (Table 5.5). Concurrently, the trityl load of 2 portions (2.15 mg, 1.90 mg) of the untreated SLCPG was also assessed, the average of which was calculated to be 53.3 μmol g^{-1}.

Table 5.5 Treatment of SLCPG with Bu^tNH_2.

Time (h)	Mass SLCPG (mg)	Vol $HClO_4$ (mL)	Abs[a]	Support load (μmol g^{-1})	Avg Support load (μmol g^{-1})
0	2.15	4.4	1.792	52.4	
0	1.90	3.8	1.896	54.2	53.3
15.7	2.24	4.4	1.960	55.0	
15.7	2.34	4.6	1.867	52.5	53.7

Support remaining following treatment with Bu^tNH_2 (%) = 53.7/53.3 * 100% = 100.8%.
[a] Absorbace values are the average of two measurements.

The difference between the loads of the treated and untreated supports was 0.8%.

5.3.9 Synthesis of sequence 5'-d(TCCTCTCT) (115)

Sequence **115** was synthesised using normal CPG on a 1 μmol scale using C^{BPA^t} phosphoramidite monomer followed by cleavage and deprotection using conc. NH_3 (aq) solution at 55 °C during 16 h. The product sequence was purified by HPLC (t_R 13.7 min) and characterised by ESMS (MW calcd 2310.4; found 2311.9). The sequence was also synthesised on SLCPG on a 1 μmol scale. The cyanoethyl and BPA^t protecting groups were removed by treating with Bu^tNH_2 and ethanolamine (Sections 5.2.4 (b) and Section 5.2.5 respectively). The fully deprotected sequence was cleaved from the support using $Et_3N.3HF$ (Section 5.2.3 (c)). The identity of the HPLC purified sample was confirmed as **115** by coinjection with an authentic sample.

5.3.10 Synthesis of the phenyl acetic acid conjugate sequence (117)

A solution of DMF (0.30 mL) containing phenylacetic acid (14.8 mg, 120 μmol), PyBOP (56.0 mg of support, 110 μmol) and DIPEA (0.05 mL, 290 μmol) was prepared and

allowed to stand for 5-10 min. After this time, the solution was added to the the deprotected SLCPG support-bound sequence **116** (0.10 μmol, 8.8 mg) in a screw-cap vial which was shaken for 2 h. After this time, the solid support was washed successively with MeCN (2 x 1 mL), DMF (2 x 1 mL), MeCN (2 x 1 mL) and Et_2O (2 x 1 mL) and then treated with 2% TFA/CH_2Cl_2 during 1 h to cleave the conjuagte sequence from the support (Section 5.2.3 (b)). Semi-preparative HPLC purification (Section 5.2.8) and analysis of the crude product mixture afforded **117** (yield not determined); HPLC: t_R 16.6 min; ESMS (-): MW calcd 2653.5 , found 2654.0.

5.3.11 Synthesis of the mitozolomide conjugate sequence (118)

Similarly, **118** was prepared using the deprotected support-bound sequence **116** (24.2 mg of support, 0.26 μmol), mitozolomide derivative **119** (25.3 mg, 100 μmol), PyBOP (52.1 mg, 100 μmol), DIPEA (0.05 mL, 290 μmol) and DMF (0.3 mL) affording **118** (yield not determined); HPLC: t_R 16.6 min; UV: λ_{max} 265.8 nm, 329.6; ESMS (-): MW calcd 2760.5 (**118**) and 2655.6 (**127**), found 2656.9. Also isolated was **109**; HPLC: t_R 13.6 min; ESMS (-): MW calcd 2578.5, found 2580.0.

5.3.12 Treatment of conjugate 118 with conc. NH_3 (aq) to give the hypoxanthine conjugate 121

A sample of **118** was treated with conc. NH_3 (aq) at 45 °C during 1 h and the crude product was analysed by UV (λ_{max} 265.6 nm). HPLC analysis and purification of the major peak afforded **121**; HPLC: t_R 14.6 min; ESMS (-): MW calcd 2655.6, found 2656.5. HPLC analysis of a mixture of **118** and **121** indicated the presence of two distinct peaks (t_R 14.4 min (**121**), 16.5 min (**118**).

5.3.13 Effect of CAP-EX voltage on the molecular ion abundance of 119 and 120 in ESMS analysis

Table 5.6 Effect of CAP-EX voltage on the stability of the tetrazinone ring of **119** and **120**.

CAP-EX (eV)	Ions present (*m/z*), abundance (%)	
	119	**120**
150 to 300	242 (0%), 137 (100%)	194 (1.7%), 137 (100%)
10	242 (0%), 137 (100%)	194 (24%), 137 (100%)

Stock solutions (28.2 μM) of **119** and **120** were prepared and analysed by ESMS (-). Two different methods were used in which different CAP-EX voltage programs were used: one

in which the CAP-EX voltage was fixed at 10 eV throughout analysis and another where the CAP-EX voltage cycled from 150 to 300 eV during analysis. The results of both methods are listed in Table 5.6.

5.3.14 Synthesis of the temozolomide conjugate 128

Similarly, **128** was prepared using the deprotected support-bound sequence (**116**) (0.10 μmol, 9.2 mg of support), temozolomide derivative (**120**) (16.8 mg, 86 μmol), PyBOP (45.0 mg, 86 μmol), DIPEA (0.05 mL, 290 μmol) and DMF (0.3 mL) (yield not determined); HPLC: t_R 14.4 min; UV: λ_{max} 264.8 nm, 329.0 nm; ESMS (-): MW calcd 2712.5 (**128**) and 2655.6 (**118**), found 2656.0.

5.3.15 Attempted synthesis of sequence 129

Sequence **129** was synthesised on a 1 μmol scale using phosphoramidite **103** (0.1 M) and SLCPG (30 μmol g^{-1}, 33.5 mg). The cyanoethyl and BPA^t protecting groups were removed by treating with Bu^tNH_2 and ethanolamine (Sections 5.2.4 (b) and Section 5.2.5) respectively. The fully deprotected support-bound sequence **131** was cleaved from the support using 2% TFA/CH_2Cl_2 during 1 h (Section 5.2.3 (b)), and purified by HPLC to afford **130** (yield not determined); HPLC: t_R 25.8 min; UV: λ_{max} 266.9 nm; ESMS (-): MW calcd 4793.9, found 4796.0.

5.3.16 Synthesis of sequence 129

Sequence **129** was synthesised as described in Section 5.3.15, but in DMT-ON mode. The last CAP cycle of DNA synthesis was omitted and the cyanoethyl and BPA^t protecting groups were removed by treating with Bu^tNH_2 and ethanolamine (Sections 5.2.4 (b) and Section 5.2.5 respectively). The DMT group was then cleaved manually using the Deblock cycle on the DNA synthesiser. The SLCPG was then washed, while still in the column, with Bu^tNH_2 (10 mL) and MeCN (20 mL) before drying by passing a stream of Ar gas through the column. The fully deprotected support-bound sequence **131** was cleaved from the support using 2% TFA/CH_2Cl_2 (Section 5.2.3 (b)), and purified by HPLC to afford **129** (6.4% yield based upon weight and load of support used); HPLC: t_R 12.7 min; UV: λ_{max} 267.2 nm; ESMS (-): MW calcd 4603.8, found 4606.5. A coupling efficiency of 60% was achieved for the addition of phosphoramidite **103** during DNA synthesis based on trityl yield.

One other major peak was present in the HPLC profile of the crude sample and was isolated affording **130**, the identity of which was confirmed by coinjection with an authentic sample: (7.5% yield); HPLC: t_R 22.2 min.

5.3.17 Synthesis of the mitozolomide conjugate sequence 132

Sequence **132** was prepared using the deprotected SLCPG support-bound sequence (**131**) (18.0 mg of support, 0.20 μmol) (Section 5.3.16). A solution of DMF (0.25 mL) containing PyBOP (2.18 mg, 4.20 μmol), mitozolomide acid (**119**) (1.02 mg, 4.20 μmol) and DIPEA (1.83 μl, 10.4 μmol) was prepared. The mixture was allowed to stand for 10 min. After this time, the yellow/brown solution was added to the solid support in a screw-cap vial which was shaken for 1 h 50 min. After this time, the solid support was washed with MeCN (2 x 1 mL), DMF (2 x 1 mL), MeCN (2 x 1 mL) and Et_2O (2 x 1 mL) and then treated with 2% TFA/CH_2Cl_2 (Section 5.2.3 (b)). HPLC analysis and purification of the crude product solution afforded **132** (4.5% yield based upon support loading); HPLC: t_R 16.5 min; UV: λ_{max} 254.5nm, 327.2 nm; ESMS (-): MW calcd 4789.8, found 4791.6 as the trisodium adduct.

Also isolated was **133**: HPLC: t_R 17.6 min; ESMS (-): MW calcd 4843.8 , found 4842.7. Other significant peaks present are listed in Table 5.7. Similarly, this reaction was carried out using 7 and 76 equivalent excesses of each of the reagents relative to the support bound oligonucleotide (**131**) (Table 5.7).

Table 5.7 HPLC analyses of the crude product mixtures obtained following reaction of **131** with varying amounts of **119**.

No. of eq of **119**	Peak retention times (t_R) (min)
7	14.0 (**129**), 16.2 (**132**), 22.5 (**130**)
21	14.5 (**129**), 16.5 (**132**), 17.6 (**133**), 18.3, 22.6 (**130**), 23.3, 24.1
76	17.9 , 19.4, 20.1, 21.1, 22.6, 24.8, 25.4, 26.1, 26.3, 26.4

5.3.18 Synthesis of the temozolomide conjugate sequence 134

Similarly, **134** was prepared in 8% yield using the deprotected support-bound sequence **131** (0.19 μmol, 16.7 mg of support), temozolomide acid (0.76 mg, 3.9 μmol), PyBOP (2.03 mg, 3.9 μmol), DIPEA (1.7 μl, 9.7 μmol) and DMF (0.25 mL); HPLC: t_R 15.1 min; UV: λ_{max} 269.0 nm, 329.1 nm; ESMS (-): MW calcd 4846.8, found 4846.0 as the trisodium adduct.

5.3.19 Attempted synthesis of mitozolomide conjugate sequence (118)

DNA synthesis of **118** was typically carried out on either a 0.2 or 1 μmol scale using phosphoramidite **135** (0.1 M) and SLCPG support (typical loading 30 $\mu mol^{-1}g$). The support-bound sequences were cleaved from the solid support and deprotected using TBAF-AcOH and desalted (Sections 5.2.3 (a) and Sections 5.2.4 (a)). HPLC purification of the crude mixture afforded **141** (9% yield) HPLC: t_R 15.7 min; ESMS (-): MW calcd 2734.5, found 2732.5.

The remaining peaks (t_R 14.8-15.3 min) were collected in one portion. ESMS of the crude mixture indicated the presence of sequence (**142**) (MW calcd 2644.5, found 2644.0).

A portion of the fully protected support-bound sequence was cleaved from the support by treatment with 2% TFA/CH_2Cl_2 (Section 5.2.3 (b)) and the crude sample was analysed by HPLC (t_R 27.0 min) and UV (λ_{max} 264.8, 323.3 nm). The sample was then treated with conc. NH_3 (aq) solution at 50 °C during 1 h and was again analysed by HPLC (t_R 15.0 min) and UV (λ_{max} 268.0 nm).

5.3.20 Synthesis of sequence 158

DNA synthesis of **158** was carried out on a 1 μmol scale using phosphoramidite (**157**) (0.10 M) and normal CPG support. The support bound sequence was cleaved and deprotected using conc. NH_3 (aq) solution (Section 5.2.2). Purification of the product mixture by HPLC afforded **158** (81% crude yield based on UV analysis); HPLC: t_R 15.0 min; ESMS (-): MW calcd 2524.5, found 2525.2.

5.3.21 Synthesis of the metronidazole conjugate sequence 183

DNA synthesis of **183** was carried out on a 1 μmol scale using phosphoramidite (**182**) (0.10 M) and normal CPG support. The support was split into two portions. One portion of the support bound sequence (17 mg of CPG, 0.42 μmol of oligonucleotide) was deprotected using a 20% Et_3N-pyridine solution (1 mL) during 2 h. The support was then washed with MeCN (2 x 1 mL) and Et_2O (2 x 1 mL) and deprotected with conc. NH_3 (aq) solution at ambient temperature during 30 min. The product was purified by HPLC affording **183**; HPLC: t_R 13.3 min; ESMS (-): MW calcd 2603.4, found 2604.2 .

5.3.22 Attempted synthesis of the metronidazole conjugate sequence 183

The support-bound oligonucleotide from Section 5.3.21 was cleaved and deprotected from the remaining portion of CPG (9.8 mg of CPG, 0.24 μmol of oligonucleotide) from the above reaction using conc. NH_3 (aq) solution at 60 °C during 16 h. The product was

purified by HPLC affording **184**; HPLC: t_R 12.2 min; ESMS (-): MW calcd 2492.4, found 2494.1. The HPLC profile of a mixture of oligonucleotide (**184**) and oligonucleotide (**183**) contained two distinct peaks.

5.3.23 Synthesis of sequences 27, 30, 31

Sequences **27**, **30** and **31** were each synthesised on 1 μmol scale and were cleaved and deprotected using conc. NH_3 (aq) solution (Section 5.2.2). Each was purified by HPLC and analysed by ESMS and HPLC (Table 5.8).

Table 5.8 HPLC and ESMS analysis of oligonucleoides (**27**), (**30**) and (**31**).

	Sequence	t_R (min)	MW Found	MW Calcd
27	5'-d(TTTCTTTTCTCTCTT)	13.1	4440.1	4438.7
30	5'-d(GGGGGAAAGAAAAGAGAGAA)	12.3	6369.0	6365.2
31	3'-d(CCCCCTTTCTTTTCTCTCTT)	13.3	5885.7	5884.0

5.3.24 General procedure for coupling of 3'-*O*-succinates 200-204 to long-chain alkylamine CPG (LCAA-CPG)[93]

To a solution of DMAP (14 mg, 120 μmol) and DhbtOH (16.7 mg, 102 μmol) in anhydrous DMF (4 mL) was added DCC (21 mg, 102 μmol) and appropriate succinate nucleoside (**200-204**, 70 μmol), each as a solution in anhydrous DMF (1 mL). The solution was allowed to stand for 5 min, after which LCAA-CPG (200 mg) was added. The mixture was shaken for 24 h. The support was washed successively with DMF (3 x 2 mL), MeOH (3 x 2 mL) and Et_2O (3 x 2 mL). The DMT loading of the supports was determined (Table 5.9).

Table 5.9 Support loadings determined following coupling of 3'-*O*-succinates (**200-204**) to LCAA-CPG.

Derivatised support	Support loading μmol g^{-1}
206	50
207	53
208	54
209	49
210	38

A stock of capping solution was prepared by dissolving acetic anhydride (0.7 mL, 7.42 mmol) and DMAP (180 mg, 1.47 mmol) in anhydrous pyridine (15 mL). Unreacted amino functions on the solid support were capped by treating the derivatised support (200 mg) to an aliquot of the capping solution (2 mL) with shaking for 1 h.[76] The support was then washed with MeOH (3 x 2 mL) and Et_2O (3 x 2 mL) and weighed.

5.3.25 Synthesis of sequences 240-247

Sequences **240-247** were each synthesised on 1 μmol scale on succinyl CPG and were cleaved and deprotected using conc. NH_3 (aq) (Section 5.2.2). The crude sequences were purified by semi-preparative HPLC and analysed by ESMS and HPLC (Table 5.10). Phosphoramidite (**236**) (0.07 M) were used in the synthesis of sequence (**245**) while phosphoramidite (**239**) (0.07 M) was used in the synthesis of sequences (**243**), (**244**) and (**247**).

Table 5.10 HPLC and ESMS analysis of sequences (**240-247**).

	Sequence	t_R/min	MW Found	MW Calculated
240	3'-d(TTTTCTTTTCCCCCCT)	13.8	4698.8	4697.8
241	5'-d(AAAAGAAAAGGGGGGA)	12.6	5061.4	5058.9
242	3'-d(AAAAGAA**A**AGGGGGGA)	12.4	5060.7	5058.9
243	3'-d(AAAAGAA**I**AGGGGGGA)	12.7	5076.9	5074.9
244	3'-d(AAAAGAA**L**AGGGGGGA)	12.6	5119.7	5115.0
245	3'-d(AAAAGAA**B**AGGGGGGA)	14.2	5168.0	5165.0
246	3'-d(AAAAGAA**G**AGGGGGGA)	12.7	nd	5074.9
247	3'-d(AAAA**I**AA**A**AGGGGGGA)	12.7	5062.8	5058.9

L = 2-*O*-allyl-A, B = 2-*O*-benzyl-A, I = isoG. nd = not determined.

Table 5.11 shows the step-wise coupling efficiencies achieved during synthesis of sequences (**245**) and (**244**) as well as the individual coupling efficiency for phosphoramidites (**236**) and (**239**).

Table 5.11 Coupling efficiency of phosphoramidites (**236**) and (**239**) during DNA synthesis.

Phosphoramidite	Coupling efficiency (%)	Step-wise yield (%)
236	98	94
239	91	94

5.3.26 Allyl group removal from 2-*O*-allyl-adenine to form in sequences (243) and (247)

I was generated from L in sequences (243) and (247) prior to cleavage and deprotection by the following treatment. To a solution of PPh_3 (28.8 mg, 109.80 mmol) and $Pd_2(dba)_3$-$CHCl_3$ (10.5 mg, 0.01 mmol) in THF (2.5 mL) was added HCO_2H (0.14 mL) and Bu^nNH_2 (0.36 mL). The mixture was shaken briefly and added to the fully protected support-bound sequence (20 mg, ~ 0.5 μmol oligonucleotide). The mixture was heated at 55 °C during 12 h. The support was washed with THF (2 x 2 mL), acetone (2 x 2 mL), diethyldithiocarbamate (0.1 M, pH 9.7, 2 x 1 mL during 15 min) and acteone (2 x 2 mL).

5.4 Thermal melting experiments

5.4.1 Preparation of UV stock solutions for thermal analysis

The following buffer solutions were prepared for UV thermal analysis:

Buffer A: $MgCl_2.6H_2O$ (100 mM), NaCl (500 mM), PIPES (100 mM) and spermine (5 mM) at pH 6.4.

Buffer B: NaCl (1.0 M), $MgCl_2$ (100 mM), TRIS.HCl (250 mM) at pH 7.2.

Buffer C: as for B but with spermine (5 mM)

5.4.2 Preparation of oligonucleotide solutions for UV thermal analysis

Appropriate aliquots of quantified oligonucleotide solutions in H_2O were added to the cuvette followed by the thermal analysis stock buffer (A, B or C) (130 μL). H_2O was then added to give a final concentration of 1.5 μM for each oligonucleotide in the sample (1300 μL final volume). A miniature magnetic stirrer was added to the cuvette and the solution was stirred during 5 min and sonicated during 30 min before UV thermal analysis.

5.4.3 UV thermal analysis

Thermal UV analysis of antisense oligonucleotides was conducted using a Varian Cary 1E UV-Visible Spectrophotometer with a 12 sample heating block having a temperature range (-10 °C to >100 °C). All samples were stirred continuously during data acquisition. At low temperatures, condensation and frosting were prevented by passing a dry stream nitrogen over the UV cuvettes.

The 1 cm UV cuvette containing the oligonucleotide/buffer samples was placed in the heating block and heated at a rate of 0.5 °C per min from 20 °C to 90 °C. It was then cooled from 90 °C to 0 °C and re-heated to 90 °C. The last two ramps were repeated three

times. Absorbance against temperature readings were recorded digitally at 0.5 °C intervals by Cary Thermal Analysis software running a Victor 433D personal computer.

5.4.4 Incubation of drug conjugates (132) and (134) and the free drugs (1) and (2) with duplex target 30.31

A solution (1 mL) containing **132**, **134**, **1** or **2** (10 μM), the duplex target **30.31** (10 μM each) and buffer A (100 μL) was heated in a screw-cap vial at 24 °C in a heating block during 3 days. After this time, all UV active material was collected in one portion by semi-preparative HPLC purification. The sample was lyophillised, suspended in H_2O and desalted using a NAP-10 column. Appropriate fractions were combined and evaporated to dryness. The sample was dissolved in a weakly basic ammonia solution and analysed by ESMS (-). No result was obtained from this analysis.

References

1. R. B. Silverman, *The Organic Chemistry of Drug Design and Drug Action*, Academic press, Inc., London, 1992, pp 244-276.
2. B. Deans, Ph.D. Thesis, Aston University, 1994.
3. I. K. Larsen, in *A Textbook of Drug Design and Development*, eds. P. Krogsgaard-Larsen, H. Bundgaard, Harwood Academic Publishers, Switzerland, 1991, pp 192-243.
4. I. Niculescu-Duvaz, I. Baracu, A. T. Balaban, *Chemistry of Antitumour Agents*, ed. D. E. V. Wilman, Blackie & Sons, Ltd., Glasgow, 1990, pp 63-130.
5. M. D. Threadgill as ref. 4, pp 187-201.
6. E. Lunt, C. G. Newton, C. Smith, G. P. Stevens, M. F. G. Stevens, C. G. Straw, R. J. A. Walsh, P. J. Warren, C. Fizames, F. Lavelle, S. P. Langdon, L. M. Vickers, *J. Med. Chem.*, 1987, **30**, 357-366.
7. M. F. G. Stevens, J. A. Hickman, R. Stone, N. W. Gibson, G. U. Baig, E. Lunt, C. G. Newton, *J. Med. Chem.*, 1984, **27**, 196-201.
8. E. S. Newlands, G. R. P. Blackledge, J. A. Slack, G. J. S. Rustin, D. B. Smith, N. S. A. Stuart, C. P. Quarterman, R. Hoffman, M. F. G. Stevens, M. H. Brampton, A. C. Gibson, *Br. J. Cancer*, 1992, **65**, 287-291.
9. Y. Wang, M. F. G. Stevens, W. T. Thomson, B. P. Shutts, *J. Chem. Soc. Perkin Trans. 1*, 1995, 2783-2787.
10. G. U. Baig, M. F. G. Stevens, *J. Chem. Soc. Perkin Trans. 1*, 1987, 665-670.
11. Y. Wang, R. T. Wheelhouse, L. Zhao, D. A. F. Langnel, M. F. G. Stevens, *J. Chem. Soc. Perkin Trans. 1*, 1998, 1669-1675.
12. P. R. Lowe, C. E. Sansom, C. H. Schwalbe, M. F. G. Stevens, A. S. Clark, *J. Med. Chem.*, 1992, **35**, 3377-3382.
13. R. T. Wheelhouse, M. F. G. Stevens, *J. Chem. Soc., Chem. Commun.*, 1993, 1177-1178.
14. B. J. Denny, R. T. Wheelhouse, M. F. G. Stevens, L. L. H. Tsang, J. A. Slack, *Biochemistry*, 1994, **33**, 9045-9051.
15. K. R. Horspool, Ph.D. Thesis, Aston University, 1988.
16. A. S. Clark, B. Deans, M. F. G. Stevens, M. J. Tisdale, R. T. Wheelhouse, B. J. Denny, J. A. Hartley, *J. Med. Chem.*, 1995, **38**, 1493-1504.
17. E. S. Newlands, M. F. G. Stevens, S. R. Wedge, R. T. Wheelhouse, C. Brock, *Cancer Treat. Rev.*, 1997, **23**, 35-61.
18. W. B. Mattes, J. A. Hartley, K. W. Kohn, D. W. Matheson, *Carcinogenesis*, 1988, **9**, 2065-2072.

19. A. S. Clark, Ph.D. Thesis, Aston University, 1991.
20. J. S. Eadie, M. Conrad, D. Toorchen, M. D. Topal, *Nature* (London), 1984, **308**, 201-203.
21. Y. Wang, M. F. G. Stevens, *J. Org. Chem.*,1997, **62**, 7288-7294.
22. B. F. Baker, P. B. Dervan, *J. Am. Chem. Soc.*, 1985, **107**, 8266-8268.
23. E. T. Kool, *Chem. Rev.*, 1997, **97**, 1473-1487.
24. T. A. Gourdie, K. K. Valu, L. Gravatt, T. J. Boritzki, B. C. Baguley, L. P. G. Wakelin, W. R. Wilson, P. D. Woodgate, W. A. Denny, *J. Med. Chem.*, 1990, **33**, 1177-1186.
25. A. M. Belikova, V. F. Zarytova, N. I. Grineva, *Tetrahedron Lett.*, 1967, 3557-3562.
26. T. Le Doan, L. Perrouault, D. Praseuth, N. Habhoub, J.-L. Decout, N. T. Thuong, J. Lhomme, C. Hélène, *Nucleic Acids Res.*, 1987, **15**, 7749-7760.
27. H. E. Moser, P. B. Dervan, *Science*, 1987, 645-650.
28. E. Uhlmann, A. Peyman, *Chem Rev.*, 1990, **90**, 544-584.
29. L. E. Xodo, G. Manzini, F. Quadrifoglio, G. A. van der Marel, J. H. van Boom, *Nucleic Acids Res.*, 1991, **19**, 5625-5631.
30. M. J. Gait, in *Oligonucleotide Synthesis - A Practical Approach*, ed. M. J. Gait, IRL Press, Oxford, 1984, pp 1-22.
31. M. Petersheim, D. H. Turner, *Biochemistry*, 1983, **22**, 256-263.
32. G. Felsenfeld, D. R. Davies, A. Rich, *J. Am. Chem. Soc.*, 1957, **79**, 2023-2024.
33. P. S. Miller, P. Bhan, C. D. Cushman, T. L. Trapane, *Biochemistry*, 1992, **31**, 6788-6793.
34. S. O. Doronina, J.-P. Behr, *Chem. Soc. Rev.*, 1997, 63-71.
35. B. Faucon, J.-L. Mergny, C. Hélène, *Nucleic Acids Res.*, 1996, **24**, 3181-3188.
36. C. R. Calladine, H. R. Drew, *Understanding DNA - The Molecule and How it Works*, Academic Press, London, 1997, 2nd Ed., pp 213-218
37. M. Brenowitz, D. F. Senear, M. A. Shea, G. K. Ackers, *Proc. Natl. Acad. Sci. USA*, 1986, **83**, 8462-8466.
38. J. A. Hartley, R. L. Souhami, *Cancer Chemotherapy*, eds. J. A. Hickman, T. R. Tritton, Blackwell Scientific Publications, 1993, pp 251-255.
39. G. E. Plum, Y.-W., Park, S. F. Singleton, P. B. Dervan, K. J. Breslauer, *Proc. Natl. Acad. Sci. USA*, 1990, **87**, 9436-9440.
40. T. J. Povsic, P. B. Dervan, *J. Am. Chem. Soc.*, 1989, **111**, 3059-3061.
41. G. Xiang, R. Bogacki, L. W. McLaughlin, *Nucleic Acids Res.*, 1996, **24**, 1963-1970.
42. *The Glen Report*, 1998, **11**, 9.

43. A. M. MacMillan, G. L. Verdine, *Tetrahedron*, 1991, **47**, 2603-2616.
44. M. P. Reddy, N. B. Hanna, F. Farooqui, *Tetrahedron Lett.*, 1994, **35**, 4311-4314.
45. J. C. Schulhof, D. Molko, R. Teoule, *Nucleic Acids Res.*, 1987, **15**, 397-416.
46. N. D. Sinha, P. Davis, N. Usman, J. Pérez, R. Hodge, J. Kremsky, R. Casale, *Biochimie*, 1993, **75**, 13-23.
47. H. Köster, K. Kulikowski, T. Liese, W. Heikens, V. Kohli, *Tetrahedron*, 1980, **37**, 363-369.
48. Glen Research, *Products for DNA Research Catalogue*, 1998, 32.
49. M. J. Gait, C. Pritchard, G. Slim, in *Oligonucleotides and Analogues - A Practical Approach*, ed. F. Eckstein, IRL Press, Oxford, 1991, pp 25-48.
50. R. Wu, N.-H. Wu, Z. Hanna, F. Georges, S. Narang, as ref. 30, pp 135-151.
51. A. M. Maxam, W. Gilbert, *Proc. Natl. Acad. Sci. USA*, 1977, **74**, 560-564.
52. A. Apffel, J. A. Chakel, S. Fischer, K. Lichenwalter, W. S. Hancock, *Anal. Chem.*, 1997, **69**, 1320-1325.
53. S. L. Beaucage, R. P. Iyer, *Tetrahedron*, 1993, **49**, 1925-1963.
54. E. S. Belousov, I. A. Afonina, M. A. Podyminogin, H. B. Gamper, M. W. Reed, R. M. Wydro, R. B. Meyer, *Nucleic Acids Res.*, 1997, **25**, 3440-3444.
55. W. H. Gmeiner, W. Luo, R. T. Pon, J. W. Lown, *Bioorg. Med. Chem. Lett.*, 1991, **1**, 487-490.
56. D. L. McMinn, T. J. Matray, M. M. Greenberg, *J. Org. Chem.*, 1997, **62**, 7074-7075.
57. S. J. Kim, M. P. Stone, C. M. Harris, T. M. Harris, *J. Am. Chem. Soc.*, 1992, **114**, 5480-5481.
58. H. Gao, R. Fathi, B. L. Gaffney, B. Goswami, P.-P. Kung, Y. Rhee, R. Jin, R. A. Jones, *J. Org. Chem.*, 1992, **57**, 6954-6959.
59. M. P. Wallis, C. H. Schwalbe, W. Fraser, *Nucleosides & Nucleotides*, 1997, **16**, 2053-2068.
60. N. Haginoya, A. Ono, Y. Nomura, Y. Ueno, A. Matsuda, *Bioconj. Chem.*, 1997, **8**, 271-280.
61. A. Roget, H. Bazin, R. Teoule, *Nucleic Acids Res.*, 1989, **17**, 7643-7651.
62. G. Wang, D. E. Bergstrom, *Tetrahedron Lett.*, 1993, **34**, 6725-6728.
63. Y. Nomura, Y. Ueno, A. Matsuda, *Nucleic Acids Res.*, 1997, **25**, 2784-2791.
64. As ref. 48, p 39.
65. B. A. Connolly, *Nucleic Acids Res.*, 1987, **15**, 3131-3139.
66. J. M. Coull, H. L. Weith, R. Bischoff, *Tetrahedron Lett.*, 1986, **27**, 3991-3994.
67. K. B. Grant, P. B. Dervan, *Biochemistry*, 1996, **35**, 12313-12319.

68. E. A. Lukhtanov, I. V. Kutyavin, V. V. Gorn, M. W. Reed, A. D. Adams, D. D. Lucas, R. B. Meyer Jr., *J. Am. Chem. Soc.*, 1997, **119**, 6214-6225.
69. J. Telser, K. A. Cruickshank, L. E. Morrisoin, T. L. Netzel, *J. Am. Chem. Soc.*, 1989, **111**, 6966-6976.
70. Q. Zheng, Y.-Z., Xu, P. F. Swann, *Nucleosides & Nucleotides,* 1995, **14**, 939-942.
71. T. M. Tarasow, D. Tinnermeier, C. Zyzniewski, *Bioconj. Chem.*, 1997, **8**, 89-93.
72. J.-P. Shaw, J. F. Milligan, S. H. Krawczyk, M. Matteucci, *J. Am. Chem. Soc.*, 1991, **113**, 7765-7766.
73. R. B. Meyer, Jr., J. C. Tabone, G. D. Hurst, T. M. Smith, H. Gamper, *J. Am. Chem. Soc.*, 1989, **111**, 8517-8519.
74. L. M. Khalimskaya, A. S. Levina, V. F. Zarytova, *Bioorg. Med. Chem. Lett.*, 1994, **4**, 1083-1088.
75. D. L. Wong, J. G. Pavlovich, N. O. Reich, *Nucleic Acids Res.*, 1998, **26**, 645-649.
76. T. Atkinson, M. Smith, as ref. 30, pp 35-81.
77. R. T. Pon, S. Yu, *Tetrahedron Lett.*, 1997, **38**, 3327-3330.
78. S. L. Woo, S. M. Menchen, S. Fung, USP 4 965 349/1990.
79. R. H. Alul, C. N. Singman, G. Zhang, R. L. Letsinger, *Nucleic Acids Res.*, 1991, **19**, 1527-1532.
80. R. T. Pon, S. Yu, *Nucleic Acids Res.*, 1997, **25**, 3629-3635.
81. N. N. Polushin, J. S. Cohen, *Nucleic Acids Res.*, 1994, **22**, 5492-5496.
82. B. A. Connolly, P. C. Newman, *Nucleic Acids Res.*, 1989, **17**, 4957-4974.
83. Y. Zhou, P. O. P. Ts'o, *Nucleic Acids Res.*, 1996, **24**, 2652-2659.
84. Y. Zhou, S. Chládek, L. J. Romano, *J. Org. Chem.*, 1994, **59**, 556-563.
85. Y. Zhou, L. J. Romano, *Biochemistry*, 1993, **32**, 14043-14052.
86. Y. Hayakawa, S. Wakabayashi, H. Kato, R. Noyori, *J. Am. Chem. Soc.*, 1990, **112**, 1691-1696.
87. T. J. Matray, D. J. Yoo, D. L. McMinn, M. M. Greenberg, *Bioconj. Chem.*, 1997, **8**, 99-102.
88. X. Zhang, B. L. Gaffney, R. A. Jones, *Nucleic Acids Res.*, 1997, **25**, 3980-3983.
89. T. J. Matray, M. M. Greenberg, *J. Am. Chem. Soc.*, 1994, **116**, 6931-6932.
90. L. Holmberg, USP 5 589 586/1996.
91. M. Kwiatkowski, M. Nilsson, U. Landegren, *Nucleic Acids Res.*, 1996, **24**, 4632-4638.
92. A. Routledge, M. P. Wallis, K. C. Ross, W. Fraser, *Bioorg. Med. Chem. Lett.*, 1995, **5**, 2059-2064.

93. A. J. Walsh, G. C. Clark, W. Fraser, *Tetrahedron Lett.*, 1997, **38**, 1651-1654.
94. A. K. Saha, M. Sardaro, C. Waychunas, D. Delecki, R. Kutny, P. Cavanaugh, A. Yawman, D. A. Upson, L. I. Kruse, *J. Org. Chem.*, 1993, **58**, 7827-7831.
95. J. F. Cormier, K. K. Ogilvie, *Nucleic Acids Res.*, 1988, **16**, 4583-4594.
96. M. A. McClinton, *Aldrichimica Acta*, 1995, **28**, 31-35.
97. L. J. McBride, C. McCollum, S. Davidson, J. W. Efcavitch, A. Andrus, S. J. Lombardi, *Biotechniques*, 1988, **6**, 362-367.
98. *Perseptive Product Bulletin*, No. 104.
99. Glen Research, *User Guide to DNA Modification*, 1996, 13.
100. Glen Research, *Products for DNA Research Catakogue*, 1995, 30.
101. W. Bannwarth, *Helv. Chim. Acta*, 1988, **71**, 1517-1527.
102. D. W. Will, G. Breipohl, D. Langner, J. Knolle, E. Uhlmann, *Tetrahedron*, 1995, **51**, 12069-12082.
103. N. D. Sinha, R. M. Cook, *Nucleic Acids Res.*, 1988, **16**, 2659-2669.
104. A. H. Krotz, P. G. Klopchin, K. L. Walker, G. S. Srivatsa, D. L. Cole, V. T. Ravikumar, *Tetrahedron Lett.*, 1997, **38**, 3875-3878.
105. *The Glen Report*, 1997, **10**, 1-2.
106. C. Vargeese, J. Carter, J. Yegge, S. Krivjansky, A. Settle, E. Kropp, K. Peterson, W. Pieken, *Nucleic Acids. Res.*, 1998, **26**, 1046-1050.
107. N. D. Sinha, J. Biernat, J. McManus, H. Köster, *Nucleic Acids Res.*, 1984, **12**, 4539-4557.
108. J. Coste, D. Le-Nguyen, B. Castro, *Tetrahedron Lett.*, 1990, **31**, 205-208.
109. O. N. Jensen, S. Kulkarni, J. V. Aldrich, D. F. Barofsky, *Nucleic Acids Res.*, 1996, **24**, 3866-3872.
110. K. R. Horspool, M. F. G. Stevens, C. G. Newton, E. Lunt, R. J. A. Walsh, B. L. Pedgrift, G. U. Baig, F. Lavelle, C. Fizames, *J. Med. Chem.*, 1990, **33**, 1393-1399.
111. J. K. Horton, M. F. G. Stevens, *J. Chem. Soc., Perkin Trans. 1*, 1981, 1433-1437.
112. A. E. Ashcroft, *Ionisation methods in organic mass spectrometry*, The Royal Society of Chemistry, Cambridge, 1997, pp 27-56.
113. J. B. Fenn, M. Mann, C. K. Meng, S. F. Wong, C. M. Whitehouse, *Science*, 1989, **246**, 64-71.
114. R. B. Cole (ed.), *Electrospray Ionization Mass Spectrometry - Fundamentals, Instrumentation and Applications*, John Wiley & Sons, Inc., New York, 1997, pp 131-2.
115. A. J. Walsh, K. C. Farrow and W. Fraser, *J. Pharm. Pharmacol.*, 1998, **50** (Supplement), 109.

116. D. L. McMinn, M. M. Greenberg, *Tetrahedron Lett.*, 1997, **38**, 3123-3126.
117. M. J. O'Donnell, N. Herbert, L. W. McLaughlin, *Bioorg. Med. Chem. Lett.*, 1994, **4**, 1001-1004.
118. R. Bischoff, J. M. Coull, F. E. Regnier, *Analytical Biochemistry*, 1987, **164**, 336-344.
119. A. J. Walsh, K. C. Ross, A. Routledge, M. P. Wallis, W. Fraser, *Pharmaceutical Sciences*, 1996, **2**, 33-38.
120. L. A. Carpino, J.-H. Tsao, H. Ringsdorf, E. Fell, G. Hettrich, *J. Chem. Soc., Chem. Commun.*, 1978, 358-359.
121. A. Rosowsky, J. E. Wright, *J. Org. Chem.*, 1983, **48**, 1539-1541.
122. C. F. Jewell, Jr., J. Brinkman, R. C. Petter, J. R. Wareing, *Tetrahedron*, 1994, **50**, 3849-3856.
123. G. A. Wiley, B. M. Rein, R. L. Hershkowitz, *Tetrahedron Lett.*, 1964, **86**, 2509-2513.
124. D. S. Kemp, F. Vellaccio, *Organic Chemistry*, Worth Publishers Inc., New York, 1980, 1246.
125. M. P. Wallis, I. D. Spiers, C. H. Schwalbe, W. Fraser, *Tetrahedron Lett.*, 1995, **36**, 3759-3762.
126. M. Freifelder, *Catalytic Hydrogenation in Organic Synthesis, Procedures and Commentary*, John Wiley & Sons, New York, 1979, pp 43-45.
127. D. I. Edwards, *J. Antimicrob. Chemother.*, 1993, **31**, 9-20.
128. T.-S. Lin, J.-H. Yang, M.-C. Liu, Z.-Y. Shen, Y.-C. Cheng, W. H. Prusoff, G. I. Birnbaum, J. Giziewicz, I. Ghazzouli, V. Brankovan, J.-S Feng, G.-D. Hsiung, *J. Med. Chem.*, 1991, **34**, 693-701.
129. W. L. Sung, *J. Chem. Soc., Chem. Commun.*, 1981, 1089.
130. M. P. Wallis, Ph.D. Thesis, Aston University, 1996.
131. T.-S. Lin, *Nucleic Acid Chemistry - part 4*, eds. L. B. Townsend, R. S. Tipson, John Wiley & Sons, Inc., 1991, 63-66.
132. S. Ikeda, I Saito, *Tetrahedron Lett.*, 1998, **39**, 5975-5978.
133. L. J. Maher III, B. Wold, P. B. Dervan, *Science*, 1989, **245**, 725-730.
134. S. Hildbrand, C. Leumann, *Agnew. Chem. Int. Engl.*, 1996, **35**, 1968-1971.
135. P. J. Bates, C. A. Laughton, T. C. Jenkins, D. C. Capaldi, P. D. Roselt, C. B. Reese, S. Niedle, *Nucleic Acids Res.*, 1996, **24**, 4176-4184.
136. J. S. Koh, P. B. Dervan, *J. Am. Chem. Soc.*, 1992, **114**, 1470-1478.
137. A. Ono, P. O. P. Ts'o, L.-S. Kan, *J. Am. Chem. Soc.*, 1991, **113**, 4032-4033.
138. A. Ono, P. O. P. Ts'o, L.-S. Kan, *J. Org. Chem.*, 1992, **57**, 3225-3230.
139. U. von Krosigk, S. A. Benner, *J. Am. Chem. Soc.*, 1995, **117**, 5361-5632.

140. J. Hunziker, E. S. Priestly, H. Brunar, P. B. Dervan, *J. Am. Chem. Soc.*, 1995, **117**, 2661-2662.
141. H. Brunar, P. B. Dervan, *Nucleic Acids Res.*, 1996, **24**, 1987-1991.
142. E. S. Priestly, P. B. Dervan, *J. Am. Chem. Soc.*, 1995, **117**, 4761-4765.
143. S. H. Krawczyk, J. F. Milligan, S. Wadwani, C. Moulds, B. C. Froehler, M. D. Matteucci, *Proc. Natl. Acad. Sci. USA*, 1992, **89**, 3761-3764.
144. P. A. Beal, P. B. Dervan, *Science*, 1991, **251**, 1360-1363.
145. A.-J. Cheng, M. W. Van Dyke, *Nucleic Acids Res.*, 1993, **21**, 5630-5635.
146. J. F. Milligan, S. H. Krawczyk, S. Wadwani, M. D. Matteucci, *Nucleic Acids Res.*, 1993, **21**, 327-333.
147. S. Agrawal, P. L. Iadarola, J. Temsamani, Q. Zhao, D. R. Shaw, *Bioorg. Med. Chem. Lett.*, 1996, **6**, 2219-2224.
148. V. Roig, R. Kurfurst, N. T. Thuong, *Tetrahedron Lett.*, 1993, **34**, 1601-1604.
149. S. D. Jayasena, B. H. Johnston, *Nucleic Acids Res.*, **20**, 5279-5288.
150. D. M. Gowers, K. R. Fox, *Nucleic Acids Res.*, 1999, **27**, 1569-1577.
151. B.-W. Zhou, C. Marchand, U. Asseline, N. T. Thuong, J.-S. Sun, T. Garestier, C. Hélène, *Bioconj. Chem.*, 1995, **6**, 516-523.
152. W. A. Greeberg, P. B. Dervan, *J. Am. Chem. Soc.*, 1995, **117**, 5016-5022.
153. S. C. Zimmerman, P. Schmitt, *J. Am. Chem. Soc.*, 1995, **117**, 10769-10770.
154. C.-Y. Huang, G. Bi, P. S. Miller, *Nucleic Acids Res.*, 1996, **24**, 2606-2613.
155. C.-Y. Huang, P. S. Miller, *J. Am. Chem. Soc.*, 1993, **115**, 10456-10457.
156. C. Y. Switzer, S. E. Moroney, S. A. Benner, *Biochemistry*, 1993, **32**, 10489-10496.
157. K. Groebke, J. Hunziker, W. Fraser, L. Peng, U. Diederichsen, K. Zimmermann, A. Holzner, C. Leumann, A. Eschenmoser, *Helv. Chim. Acta*, 1998, **81**, 375-474.
158. Z. Kazimierczuk, R. Mertens, W. Kawczynski, F. Seela, *Helv. Chim. Acta*, 1991, **74**, 1742-1748.
159. C. Roberts, R. Bandaru, C. Switzer, *Tetrahedron Lett.*, 1995, **36**, 3601-3604.
160. F. Seela, B. Gabler, *Helv. Chim. Acta*, 1994, **77**, 622-630.
161. V. L. Worthington, W. Fraser, C. H. Schwalbe, *Carbohydrate Res.*, 1995, **275**, 275-284.
162. Z. Kazimjerczuk, H. B. Cottam, G. R. Revankar, R. K. Robins, *J. Am. Chem. Soc.*, 1984, **106**, 6379-6382.
163. M. Hoffer, *Chem. Ber.*, 1960, **93**, 2777-2781.
164. T. Tuschl, M. M. P. Ng, W. Pieken, F. Benseler, F. Eckstein, *Biochemistry*, 1993, **32**, 11658-11668.

165. M. M. P. Ng, F. Benseler, T. Tuschl, F. Eckstein, *Biochemistry*, 1994, **33**, 12119-12126.
166. R. A. Jones, as ref. 30, pp 29-30.
167. F. Seela, C. Wei, Z. Kazimierczuk, *Helv. Chim. Acta*, 1995, **78**, 1843-1854.
168. F. Seela, C. Wei, *Helv. Chim. Acta*, 1997, **80**, 73-85.
169. L. J. McBride, R. Kierzek, S. L. Beaucage, M. H. Caruthers, *J. Am. Chem. Soc.*, 1986, **108**, 2040-2048.
170. F. Seela, C. Wei, A. Melenewski, *Nucleic Acids Res.*, 1996, **24**, 4940-4945.
171. T. Horn, C.-A. Chang, M. L. Collins, *Nucleosides & Nucleotides,* 1995, **14**, 1023-1026.
172. M. J. Lutz, J. Horlacher, S. A. Benner, *Bioorg. Med. Chem. Lett.*, 1998, **8**, 499-504.
173. *Aldrich Catalogue*, 1999-2000, 1141.
174. T. Brown, D. J. S. Brown, as ref. 49, pp 1-23.
175. Amersham Pharmacia Biotech, *Instructions for Using NAP 10 Columns*.