Aston University

Some pages of this thesis may have been removed for copyright restrictions.

If you have discovered material in AURA which is unlawful e.g. breaches copyright, (either yours or that of a third party) or any other law, including but not limited to those relating to patent, trademark, confidentiality, data protection, obscenity, defamation, libel, then please read our Takedown Policy and contact the service immediately

Nucleoside and Oligonucleotide Approaches Towards Potential HIV Chemotherapies

Mark Philip Wallis

Doctor of Philosophy

THE UNIVERSITY OF ASTON IN BIRMINGHAM

SEPTEMBER 1996

This copy of the thesis has been supplied on the condition that anyone who consults it is understood to recognise that its copyright rests with the author and that no quotation from the thesis and no information derived from the it may be published without proper acknowledgement.

THE UNIVERSITY OF ASTON IN BIRMINGHAM

Nucleoside and Oligonucleotide Approaches Towards Potential HIV Chemotherapies

A thesis submitted by Mark Philip Wallis BPharm (Hons) for the degree of

Doctor of Philosophy

Abstract: The use of oligonucleotides directed against the mRNA of HIV promises site-specific inhibition of viral replication. In this work, the effect of aralkyl substituents on oligonucleotide duplex stability was studied using model oligonucleotide sequences in an attempt to improve targeting of oligonucleotides to viral mRNA.

Aralkyl-substituted oligonucleotides were made by solid phase synthesis using either the appropriate aralkyl-substituted phosphoramidite or by post-synthetic substitution of a pentafluorophenoxy substituent by *N*-methylphenethylamine. The presence of phenethyl or benzoyl substituents invariably resulted in thermodynamic destabilisation of all duplexes studied.

The methods which were developed for the synthesis of nucleoside intermediates for oligonucleotide applications were also used to prepare a series of nucleoside analogues derived from uridine, 2'-deoxyuridine and AZT. Crystal structures of six compounds were successfully determined. Anti-HIV activity was observed for most compounds in the series although none were without cytotoxicity. The most active compound of the series was the ribose nucleoside; 1-β-D-*erythro*-pentofuranosyl-4-pentafluorophenoxy-pyrimidine-2(1*H*)-one **95**, derived directly from uridine. The same series of compounds also displayed very modest anti-cancer activity.

To enable synthesis of prooligonucleotides and analogues for possible antisense applications, the properties of a new Silyl-Linked Controlled Pore Glass solid support were investigated. Synthesis of the sequences $d(Tp)_7T$, $d(Tps)_7T$ and the base-sensitive $d(Tp)_3(C^{Bz}p)_2(Tp)_2T$ was achieved using the silyl-linked solid support in a fluoride-induced cleavage/deprotection strategy.

Keywords: Anti-HIV activity, Nucleoside, Antisense oligonucleotide, Crystal structure, Fluoride cleavage/deprotection.

I would very much like to thank my supervisor Dr. William Fraser for always being approachable, his reassuring guidance and his constant patience during my time at Aston.

I would also like to thank Dr. Anne Routledge for the supply of silyl-linked solid support for investigational Dr. Carl Bolton and Dr. Jon Hobbs for their help in crystal structure determination.

Thanks also to Nahwed Mahmood (MRC Collaborative Centre, Cambridge), Augustin Dick and Stella M. Mike J. Tisdale (Aston) for biological test results.

Thanks also to the BBSRC for the financial support for this project.

Thanks to all my friends in the University, especially my badminton partner Karen, for making my stay so memorable.

Finally, thanks to my friends outside of Aston, Simon and Nigel who both helped in their uniquely individual ways.

To my family and Simon, your support did not go unnoticed.

Acknowledgements

I would very much like to thank my supervisor Dr. William Fraser for always being approachable, his reassuring guidance and his constant patience during my time at Aston.

I would also like to thank Dr. Anne Routledge for the supply of silyl-linked solid support for investigation and Dr. Carl Schwalbe and Dr. Ian Spiers for their help in crystal structure determination.

Thanks also to Naheed Mahmood (MRC Collaborative Centre, Cambridge), Augustin Dick and Professor Mike J. Tisdale (Aston) for biological test results.

Thanks also to the BBRSC for the financial support for this project.

Thanks to all my friends in the University, especially my badminton partner Karen, for making my stay so memorable.

Finally, thanks to by friends outside of Aston, Simon and Nigel who both helped in their uniquely individual ways.

Contents

Chapter 1 HIV: Current and Potential Strategies Towards its Demise

Chapter 5 A New Cleavage/Deprotection Strategy for Automated Oligonucleotide Synthesis Using a Novel Silyl-Linked Solid Support

Chapter 6 Experimental

List of Figures

List of Tables

Chapter 1

HIV: Current and Potential Strategies Towards its Demise

1.1 Introduction

The Human Immunodeficiency Virus (HIV) is the virulent agent responsible for Acquired Immunodeficiency Syndrome (AIDS). Generally, HIV infection is fatal with no known cure.[1] However, individuals who have shown signs of exposure to HIV but remain persistently seronegative, have been identified.[2] The World Health Organisation estimates that by the turn of the century in excess of 20 million individuals globally will be infected with HIV[3]. To date, health education remains the only means of slowing the spread of HIV infection throughout the human population.[4,5]

1.2 The Human Immunodeficiency Virus

HIV is classed as a Lentivirus. Many different isolates of HIV have been detected which have been grouped broadly as being HIV-1 or HIV-2.[6] The HIV viron contains a ribonucleoprotein core particle surrounded by an outer protein shell (Figure 1.1)[7].

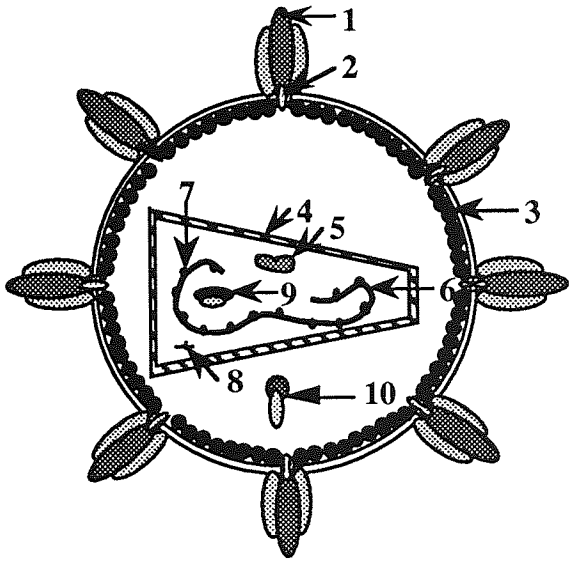


Figure 1.1 The HIV viron. 1 = surface glycoprotein (gp 120); 2 = transmembrane glycoprotein (gp 41); 3 = matrix protein (p 17); 4 = capsid protein; 5 = integrase; 6 = genomic RNA; 7 = nucleocapsid protein; 8 = tRNA; 9 = reverse transcriptase; 10 = protease.

The diameter of the quasispherical particle is about 10 nm. The 'clubbed' shaped inner core consists of capsid proteins which are associated with two copies of the (+)-stranded HIV genomic RNA[8]. Within the core are molecules of reverse transcriptase and integrase. The core is surrounded by a matrix protein which in turn is surrounded by a lipid bilayer studded by glycoprotein (gp 120) in association with the transmembrane hydrophobic glycoprotein (gp 41).

1.2.1 Infectious cycle

The infectious cycle of HIV is summarised in Figure 1.2. Infection of a new host cell starts with gp 120 surface glycoproteins of HIV interacting with CD4 receptors of the host. Fusion of the viral and cellular membranes follows, in which the transmembrane glycoprotein plays a central role, to release the subviral particle into the host cell.[9]

Reverse transcription of (+)-stranded HIV genomic ribonucleic acid (RNA) into the double-stranded deoxyribonucleic acid (DNA) 'pro-virus' by viral reverse transcriptase (RT) then follows. Reverse transcription is a process of retroviruses that significantly differs from the biochemistry of the host cell, providing opportunities for anti-viral therapies. RT polymerises DNA using the viral RNA as a reverse transcript to form a RNA·DNA heteropolymer. This process is known as RNA-dependant DNA polymerisation (RDDP). Reverse transcriptase has ribonuclease H (RNase H) activity which degrades the original RNA genome. DNA-dependant DNA polymerisation (DDDP) by reverse transcriptase using the DNA transcript of the viral RNA genome yields the pro-viral DNA·DNA homopolymer. The complete pro-viral double-stranded DNA is integrated into the host cell chromosome by viral integrase. The host cell, under partial control of viral regulatory proteins, transcribes viral genes and translates the resulting messenger RNA (mRNA) to make regulatory and structural viral proteins. The viral genome is transcribed by the host cellular machinery and is assembled with structural and unprocessed metabolic proteins. The immature virus then buds from the cell and maturation follows.

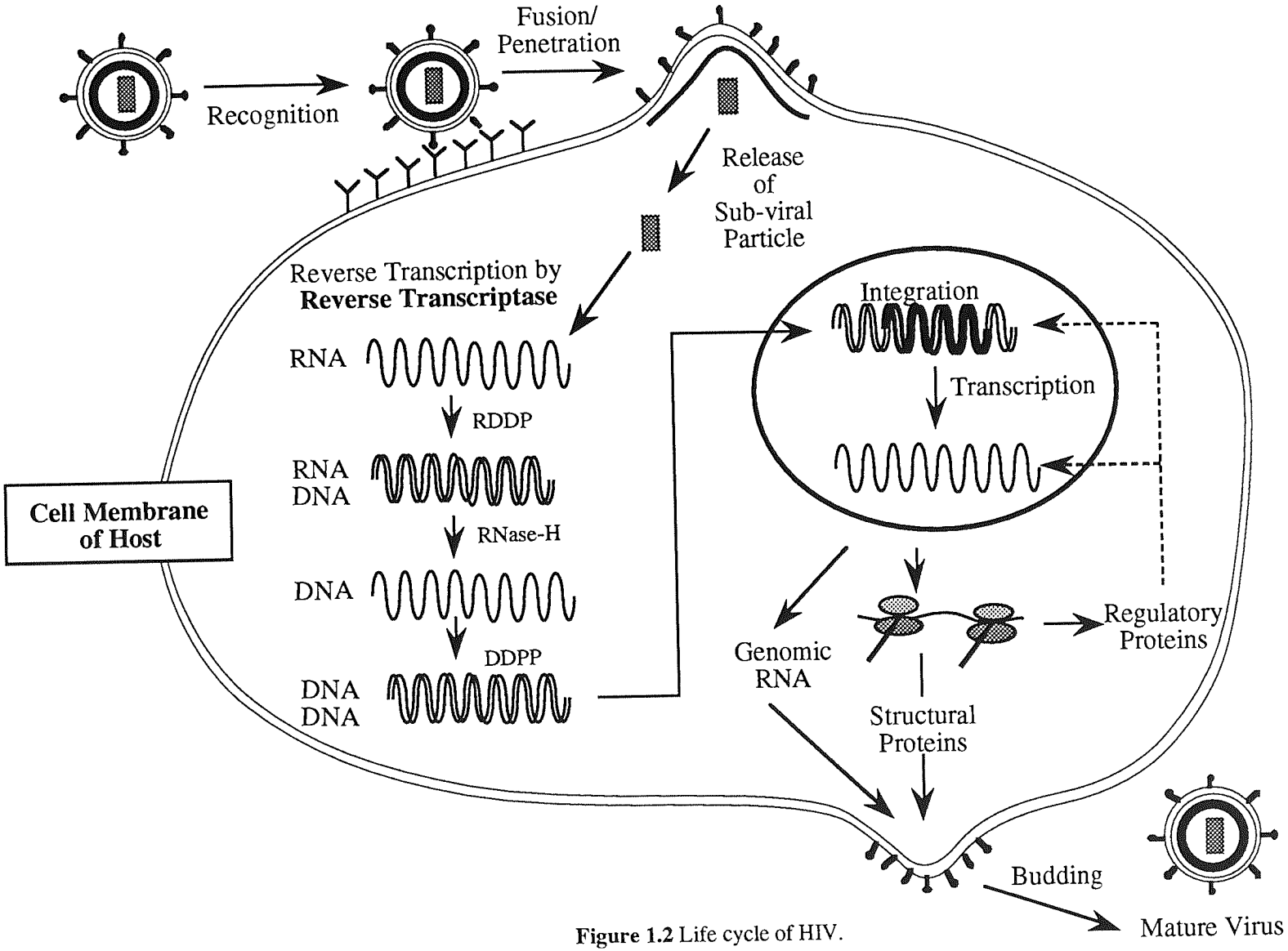


Figure 1.2 Life cycle of HIV.

1.2.2 Genetic structure of HIV

HIV genomic RNA contains the genetic information of three major genes: *gag*, *pol* and *env*, which make up the bulk of the HIV viron and regulatory proteins. In addition, HIV encodes for a number of regulatory proteins of which rev and tat are essential for viral replication and provide valuable targets for anti-HIV therapy.[10]

1.2.3 Regulatory proteins

Viral regulatory proteins influence replication of HIV by influencing the cellular machinery of the host. The HIV-1 gene codes for regulatory proteins: tat, rev, nef, vif, vpu and vpr (Table 1.1). The HIV-2 gene codes for analogues of the above proteins plus the additional vpx regulatory protein. In terms of drug design directed against HIV, the rev and tat regulatory proteins offer potential targets for modulating HIV replication.

Product	Function
tat	Protein transcription factor
rev	Regulator of gene expression
vpr	Weak transcriptional factor
vpu	Required for efficient virus budding
vif	Promotes infectivity of free virus
nef	Viral pathogenesis
LTR	Binding site for viral and host transcription factor

Table 1.1 Protein products of the HIV genome.

The trans-activator of transcription (tat) regulatory protein is critical for HIV gene expression.[11] Increase in levels of viral gene expression brought about by tat result from interactions between tat and the trans-activation responsive region (TAR) of the HIV long terminal repeats (LTR). Antisense oligonucleotides directed against TAR could potentially inhibit tat-TAR interactions, blocking or slowing viral replication.

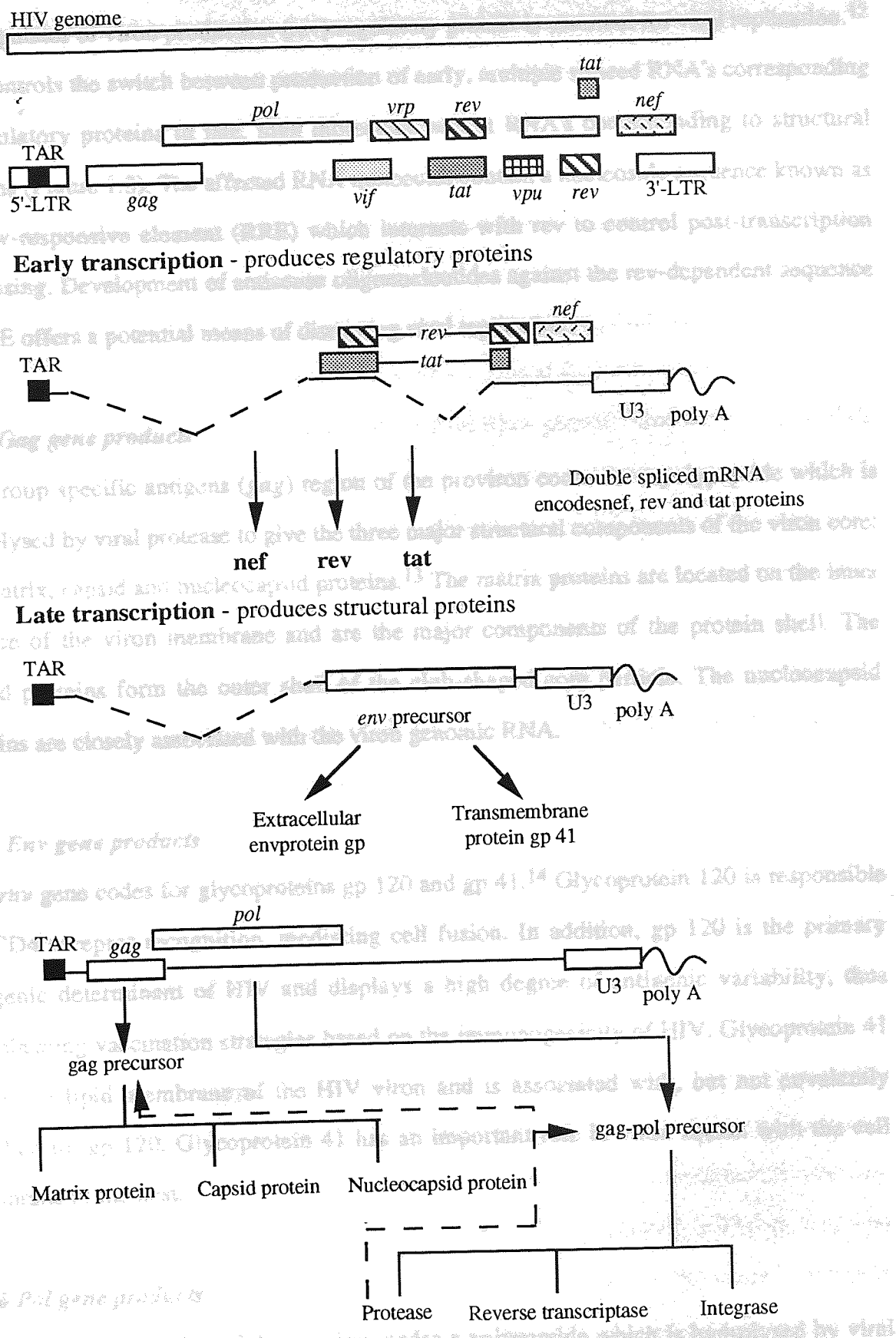


Figure 1.3 The HIV genome and its protein products.

The regulator of viron production (rev) regulatory protein is essential for viral replication.[12] Rev controls the switch between production of early, multiple spliced RNA's corresponding to regulatory proteins to late, high molecular weight RNA's corresponding to structural proteins (Figure 1.3). The affected RNA molecules contain a nucleoside sequence known as the rev-responsive element (RRE) which interacts with rev to control post-transcription processing. Development of antisense oligonucleotides against the rev-dependent sequence of RRE offers a potential means of disrupting viral replication.

1.2.4 Gag gene products

The group specific antigens (*gag*) region of the proviron codes for a polypeptide which is hydrolysed by viral protease to give the three major structural components of the viron core: the matrix, capsid and nucleocapsid proteins.[13] The matrix proteins are located on the inner surface of the viron membrane and are the major components of the protein shell. The capsid proteins form the outer shell of the club-shaped core particle. The nucleocapsid proteins are closely associated with the viron genomic RNA.

1.2.5 Env gene products

The *env* gene codes for glycoproteins gp 120 and gp 41.[14] Glycoprotein 120 is responsible for CD4 receptor recognition, mediating cell fusion. In addition, gp 120 is the primary antigenic determinant of HIV and displays a high degree of antigenic variability, thus complicating vaccination strategies based on the immunogenicity of HIV. Glycoprotein 41 spans the lipid membrane of the HIV viron and is associated with, but not covalently attached to, gp 120. Glycoprotein 41 has an important role in viron fusion with the cell membrane of the host.

1.2.6 Pol gene products

Polymerase (*pol*) region of the proviron codes a polypeptide which is hydrolysed by viral protease to give viral protease, integrase and reverse transcriptase.

Protease is responsible for proteolytic processing of HIV *gag* and *gag-pol* polyproteins.[15] Viral protease provides a unique target for HIV therapy; anti-HIV protease inhibitors are discussed in Section 1.4.2.

Reverse transcriptase is responsible for reverse transcription of genomic HIV RNA to the pro-virus.[16] RT has a relatively high rate of misincorporation.[17] Unlike cellular DNA polymerases, HIV reverse transcriptase has no 3'-5' exonuclease proof-reading mechanism. The most common errors are base shift mutations caused by primer-template slippage. Errors can also occur during transcription of the viral RNA genome from the pro-viral DNA by host cell RNA polymerase II. RNase H activity of reverse transcriptase is responsible for digestion of the RNA template during reverse transcription.[18] RT is thought to have both endo and exonuclease activity. The endonucleolytic activity of RT provides a suitable target for antisense oligonucleotides.

Integrase is responsible for incorporating the pro-virus into the host's genome but does not appear to be essential for viral replication.[19]

1.3 From HIV Infection to AIDS

After initial infection, the critical event is localisation of the HIV virus in the lymphoid organs which act as reservoirs for HIV infection and HIV replication[20] although, at any given time, HIV can be detected in the plasma. HIV can spread through lymphoid tissue cell-to-cell. There is a vigorous humoral (antibodies against HIV proteins) and cell-mediated (HIV-specific T-lymphocytes) immune response to HIV.[21-24] B-Lymphocytes are associated with antibody production and T-lymphocytes are associated with cell-mediated immunity. The initial immune response to HIV infection is plotted in Figure 1.4. The measurement of p24 HIV antigen in serum is a measure of viral load but is not necessarily associated with intact viral particles. Viron-associated RNA is found in plasma at high concentrations during initial, acute HIV infection but decreases 1000-fold at seroconversion. HIV RNA, however, is still detectable during clinical latency.[25]

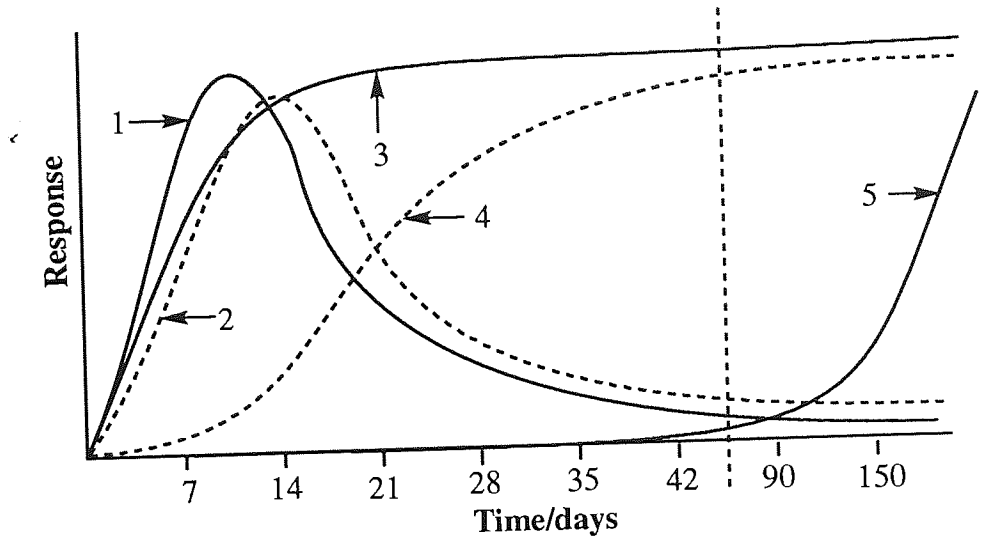


Figure 1.4 1 = virus-expressing cells in lymphatic tissue; 2 = p 24; 3 = virus-specific cytotoxic T lymphocytes; 4 = virus trapping; 5 = neutralising antibody.

CD4+ lymphocytes, the so called helper cells, interact with other components of the immune system through the production of cytokines. CD8+ lymphocytes have two functions; some act as specific killer cytotoxic T-cells which destroy HIV infected host cells, whilst others suppress the immune response once an infection has been overcome.

Initially there is a strong CD8+ lymphocyte response. The appearance of virus-specific cytotoxic T-lymphocytes causes a rapid decrease in the number of virus-expressing host cells. This cell-mediated response succeeds in dramatically reducing the HIV viral load but, for a number of possible reasons, fails to eradicate the infection.[26] About 90 days after infection, HIV-neutralising antibodies can be detected at which point seroconversion is said to have occurred. A 'clinically latent' phase which can last for over 10 years follows.

HIV enters CD4+ lymphocytes *via* its CD4 receptors. It has been estimated that between 1:100 and 1:1000 CD4+ lymphocytes are infected at any one time. There is a fall in CD4+ lymphocyte count and defects in their function can be detected, which predicts the development of AIDS. Specific CD8+ function also fails. Although the number of monocytes/macrophages is normal in HIV patients, their function appears to be

compromised. These observations are associated with the production of cytokines of which $CD4^+$ lymphocyte produce a number. Macrophages, which also have $CD4^+$ receptors, can be infected with HIV affecting their function. Macrophage infection also stimulates the production of a variety of cytokines. After initial infection, $CD4^+$ lymphocyte count falls leading to a reversal in the $CD4^+$:$CD8^+$ ratio.[27]

Recent research suggests that the replication rate of HIV is extremely high leading to a high death rate of $CD4^+$ cells; the rate of $CD4^+$ cell replacement is also high.[28,29] It has been suggested that $CD8^+$ cells destroy virally infected $CD4^+$ lymphocytes in the region of billions per day. The battle is thought to be evenly balanced, appearing as a 'clinically latent' phase. It is possible that such extraordinary high levels of activity cannot be maintained and eventually the immune system is overwhelmed.

The over-running of the immune system by the sheer magnitude of HIV replication appears to be over simplified. Disorientation of immune systems mediated by cytokines is thought to be critical for the progression from HIV infection to AIDS. Cytokines are proteins produced by cells of the immune system to act on cells of the immune system; the so called cytokine network. There is growing evidence that HIV infection disrupts the normal balance of cytokines. Cytokine imbalance may help HIV target $CD4^+$ cells and lymphoid tissue.

The hypothesis that cytokine imbalance causes the progression from HIV infection to AIDS is called the Th-1/Th-2 theory and has been extensively reviewed.[30,31] Th-1 refers to cytokines produced by T-lymphocytes in response to HIV infection and include γ-interferon and interleukins 2 and 12 (IL-2, IL-12) which are associated with cellular immune responses such as $CD8^+$ lymphocyte response. During asymptomatic periods of HIV infection, the level of these cytokines remains high, which is thought to suppress the levels of Th-2 cytokines. For individuals who progress from HIV infection to AIDS, a shift in the balance of Th-1/Th2 occurs. The levels of Th-1 cytokines falls with a rise in the levels of Th-2 cytokines which include IL4, IL5, IL6, IL10 and tumour necrosis factor (TNF)-γ. Th-2

cytokines are associated with the humoral immune responses. It is the reduced activity of $CD8^+$ cytotoxic T-cells which is thought to allow the balance to tip in the favour of HIV.[32,33] Figure 1.5 summarises the progression from HIV infection to AIDS.

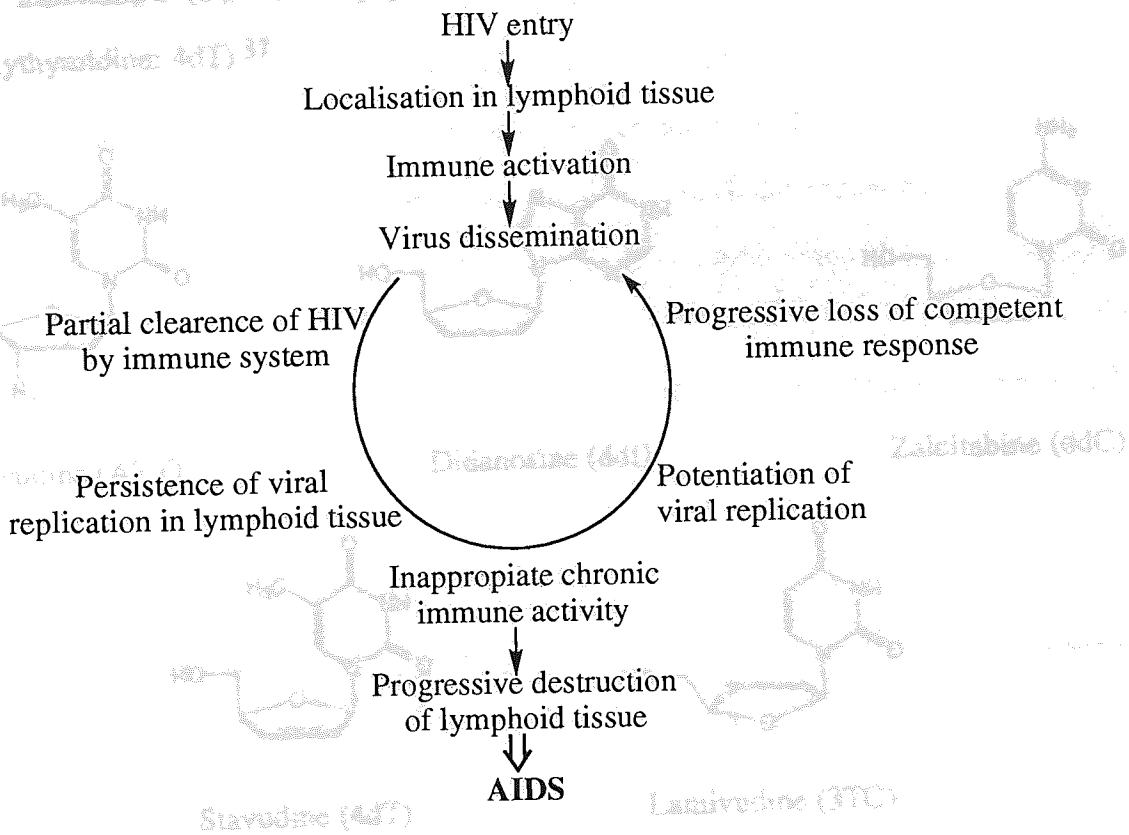


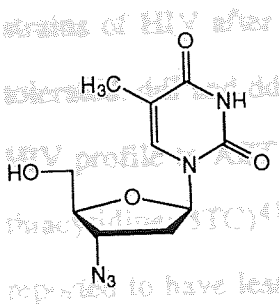
Figure 1.5 Summary of the progression from HIV infection to AIDS.

1.4 Therapeutic Strategies

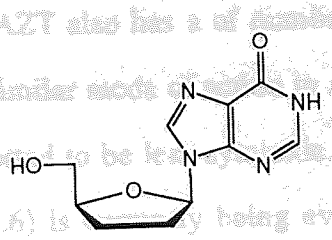
Anti-HIV drug research is divided into two broad categories: immunological and non-immunological. Immunology includes the development of vaccines (Section 1.4.4) and the potential use of immunological mediators to produce a competent immune response. Non-immunological drugs include natural products, synthetic molecules and antisense oligonucleotide technologies and such approaches to anti-HIV therapy are discussed in more detail below.

1.4.1 Reverse transcriptase inhibitors

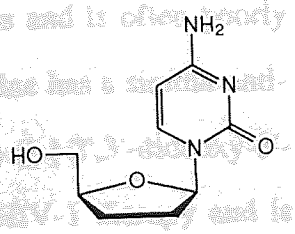
Anti-viral nucleoside analogues remain the corner-stone of HIV therapy. Around the world, four anti-viral nucleoside analogues have been licensed for the treatment of HIV infection (Figure 1.6): Zidovudine (3'-azidothymidine; AZT),[34] Didanosine (2',3'-dideoxyinosine; ddI),[35] Zalcitabine (2',3'-dideoxycytidine; ddC)[36] and Stavudine (2',3'-didehydro-2',3'-dideoxythymidine; 4dT).[37]



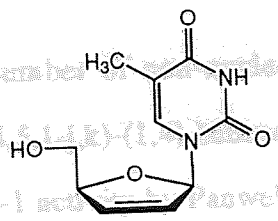
Zidovudine (AZT)



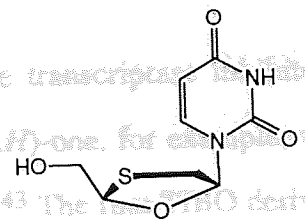
Didanosine (ddI)



Zalcitabine (ddC)



Stavudine (4dT)



Lamivudine (3TC)

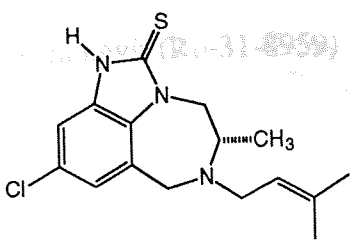
Figure 1.6 Nucleoside inhibitors of reverse transcriptase.

AZT was first synthesised in 1964[38] and found to have anti-viral activity in 1974,[39] its anti-HIV-1 activity was discovered in 1985.[40] AZT is a pro-drug which is phosphorylated by intracellular kinases to give the biologically active nucleoside 5'-triphosphate. HIV does not code for any kinase enzymes nor does it induce kinase activity in the cell, therefore, the rate of phosphorylation is dependent on host cell type. Phosphorylated AZT is thought to inhibit the incorporation of thymine by viral RT into viral DNA. Once incorporated into viral DNA, the 3'-azido group acts as a 'full stop', preventing further polymerisation of the viral DNA, thus causing chain termination. RT is 300 times more selective for AZT than cellular

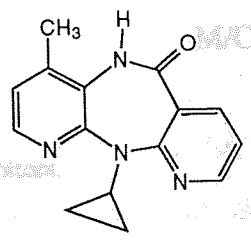
DNA polymerases while mitochondrial DNA polymerase γ is unaffected. DNA repair mechanisms of the host cell do not appear to repair viral DNA.

By the nature of its mechanism, AZT can only inhibit the infection of uninfected cells. Production of viral particles by infected cells is unaffected. Resistance to AZT by the HIV virus develops. Patients with advanced AIDS can be expected to have AZT-resistant HIV strains of HIV after one year. AZT also has a of number side-effects and is often poorly tolerated. ddI and ddC share a similar mode of action to AZT. Stavudine has a similar anti-HIV profile to AZT but is reported to be less cytotoxic. Lamivudine ([-]-2',3'-dideoxy-3'-thiacytidine; 3TC)[41] (Figure 1.6) is currently being evaluated for HIV-1 therapy and is reported to have less potential to cause bone marrow suppression or peripheral neuropathy compared with other nucleoside inhibitors of RT.[42]

There exists a number of non-nucleoside reverse transcriptase inhibitors. TIBO (tetra-hydro-imidazo-(4,5,1-j,k)-(1,4)-benzodiazepin-2-(1*H*)-one, for example, was demonstrated to have anti HIV-1 activity by Pauwels et al 1990.[43] The first TIBO derivative to be tested in patients was 9-chloro-TIBO which is a non-competitive inhibitor of RT with respect to 2'-deoxynucleotide-5'-triphosphates and an non-competitive inhibitor with respect to the primer/template of HIV-1 reverse transcriptase (Figure 1.7).[44] Resistance to TIBO and its derivatives develops rapidly.[45]



9-Chloro-TIBO



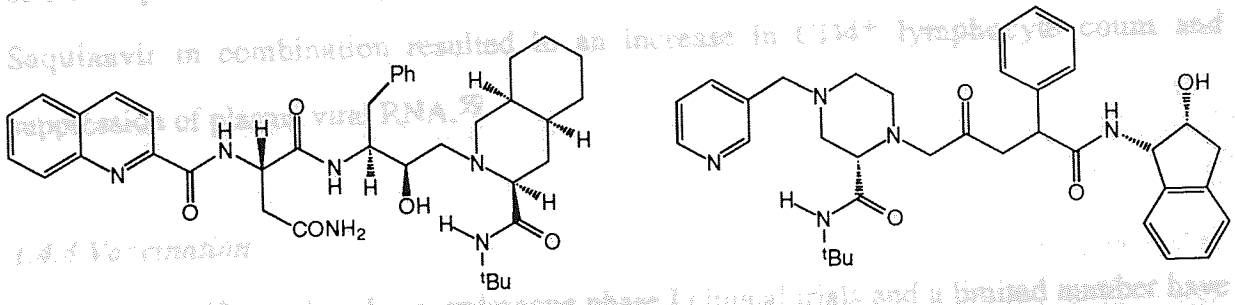
Nevirapine

Figure 1.7 Non-nucleoside inhibitors of reverse transcriptase.

Nevirapine is a dipyridodiazepinone (Figure 1.7), acting as a non-competitive inhibitor of reverse transcriptase with respect to 2'-deoxyguanosine-5'-triphosphate (dGTP).[46] Nevirapine is synergistic with AZT but inactive against TIBO-resistant HIV strains. Like TIBO, HIV resistance towards nevirapine develops rapidly.[47]

1.4.2 Protease inhibitors

HIV protease is essential for HIV infectivity,[48] cleaving the products of the *gag* and *gag-pol* genes. Saquinavir (Ro-31-8959)[49] is a protease inhibitor currently undergoing phase III clinical trials (Figure 1.8).[50] Saquinavir has potent activity against HIV protease but has poor oral bioavailability and requires frequent large dosing. HIV develops rapid resistance to Saquinavir over a relatively short period of time.[51] M/C-639 (MK-639; L-738,524)[52] is a Merck protease inhibitor currently completing phase II trials (Figure 1.8).[50] Like Saquinavir, HIV resistance to M/C-639 develops quickly.[53] After six weeks of taking M/C-639, some patients show a rebound in plasma viral RNA but p 24 remains suppressed. $CD4^+$ lymphocyte count and clinical benefits such as weight gain can still be seen after 24 weeks.



Saquinavir (Ro-31-8959) M/C-639

Figure 1.8 Protease inhibitors.

1.4.3 Combination therapy

The use of more than one anti-HIV therapy in combination has a number of potential advantages:[54, 55]

i. greater suppression of viral replication since evidence suggests that combination therapy leads to better maintenance and restoration of the immune system;
ii. inhibition or delay of appearance of drug-resistant HIV strains and
iii. possible targeting of cellular or tissue reservoirs.

To date, all combination therapies have consisted of reverse transcriptase inhibitors with other reverse transcriptase inhibitors and/or protease inhibitors. Studies have concluded that combined AZT and ddI therapy provided a sustained elevation in $CD4^+$ lymphocytes and was superior to alternating AZT and ddI therapies.[56] Alternating therapies of AZT and ddC reduced toxicity associated with each drug alone, sustaining a decrease of p 24 antigen during the trial. Patients receiving a combination therapy of AZT and ddC showed reduced disease progression and death rate compared to those receiving AZT alone.[57] Combination of Nevirapine and AZT did not prolong Nevirapine activity.[58] Trials of AZT, ddI and Saquinavir in combination resulted in an increase in $CD4^+$ lymphocyte count and suppression of plasma viral RNA.[59]

1.4.4 Vaccination

To date, over 17 vaccines have undergone phase I clinical trials and a limited number have completed phase II trials.[60] Authority to conduct large scale phase III trials has been denied or delayed as experimental vaccines have so far failed confer protection in various animal and model systems. HIV vaccines have failed to elicit an immune response capable of neutralising HIV-1 in human candidates. For this very reason, the U.S. Institute for Health recently rejected large scale phase III trials of recombinant gp120 vaccine.[61]

Despite the uncertainty of providing absolute protection against HIV, future vaccination strategies may be able to limit infection and induce an appropriate immune response that prevents disease and transmission.[62, 63]

1.5 Oligonucleotides

The term antisense oligonucleotide refers to a short polymer of nucleosides, connected *via* a linker. Oligonucleotides are found endogenously in eucaryotic cells[64] and development of automated solid-phase nucleic acid chemistry[65] has allowed their rapid synthesis in the laboratory. Endogenous oligonucleotides have phosphodiester linkage connecting the 3'-hydroxyl of one nucleoside to the primary 5'-hydoxyl of the adjacent nucleoside.

Antisense oligonucleotides hybridise to complementary sequences of target nucleic acid, generally in a Watson-Crick manner[66], to effect a response. Gene expression in cells is controlled by DNA-binding proteins, repressors and activators. Endogenous ribonucleic acids (RNA) have been shown to act as repressors in regulating gene function by blocking gene expression. Activation of gene expression by antisense oligonucleotides is achieved by suppressing the biosynthesis of repressors or by inhibiting termination of transcription.

1.5.1 Modified antisense oligonucleotides

2'-Deoxyribonucleic acid (DNA) and RNA oligonucleotides are readily accessible using automated solid-phase synthetic chemistry which has also proved versatile for synthesising oligonucleotide analogues.[65,67,68] Novel nucleosides with modified sugars or heterocycles, for example, can be introduced site-specifically into an antisense oligonucleotide. The introduction of 'non-natural' phosphorothioate and methylphosphonate backbones has received much attention in the literature as they are poor substrates for nucleases unlike phosphodiester-linked oligonucleotides which are rapidly degraded. Automated oligonucleotide synthesis usually requires the use of concentrated aqueous ammonia to fully deprotect the oligonucleotide. Such a deprotection strategy precludes the synthesis of oligonucleotides sensitive to aqueous ammonia. Chapter 5 discusses cellular uptake,

chemical stability and effectiveness of phosphodiester, phosphorothioate and methylphosphonate antisense oligonucleotides and the evaluation of a new strategy allowing the synthesis of oligonucleotides sensitive to aqueous ammonia deprotection.

1.5.2 Antisense oligonucleotides as anti-HIV therapeutic agents

Knowledge of the sequence of the intended nucleic acid target or 'receptor' allows the synthesis of an antisense oligonucleotide inhibitor or 'antagonist', which amounts to rational drug design.[69] The detailed mechanisms of antisense oligonucleotides are discussed in Section 1.8. In terms of anti-HIV therapy, antisense oligonucleotides are thought to be less likely to induce resistant strains of HIV as 'immutable' sequences of the HIV genome can be targeted; sites which are difficult to target at the protein level. The use of combinations of antisense oligonucleotides against discrete sites of the target nucleic acid should also reduce the emergence of drug resistant HIV strains. However, a number of problems are associated with the use of antisense oligonucleotides as anti-HIV agents. Phosphodiester oligonucleotides have a short half-life due to hydrolysis by plasma nucleases and have poor cellular uptake (Chapter 5). A number of synthetic DNA[70-78] and RNA[79-82] antisense oligonucleotides have been directed against HIV, targeted towards *gag*[70,71,70] *tat*[72-74,79,81] and *rev*[76,79] genes; splicing sequences,[77] tRNA primer sites[70,82] and the poly-A cap[72] using phosphodiester,[70,72,73,79-82] phosphorothioate[71,75,76,78] and methylphosphonate backbones.[73,74] Ribozymes directed against HIV have also been described.[83,84]

In order to enhance the effectiveness of antisense oligonucleotides it is important to consider:

i. nucleic acid structure;

ii. the nature of the interaction between hybridised duplex nucleic acids and

iii. potential targets for antisense oligonucleotides.

1.6 The Structure of Nucleic Acids

Oligonucleotides generally hybridise to their nucleic acid targets according to Watson-Crick hydrogen bonding rules due to mutual recognition of T by A and C by G[66,85] (Figure 1.9). The affinity and specificity of oligonucleotides for their nucleic acid targets is discussed in Section 1.7.1 and the stabilising forces of nucleic acid duplexes are discussed in Section 1.7.2.

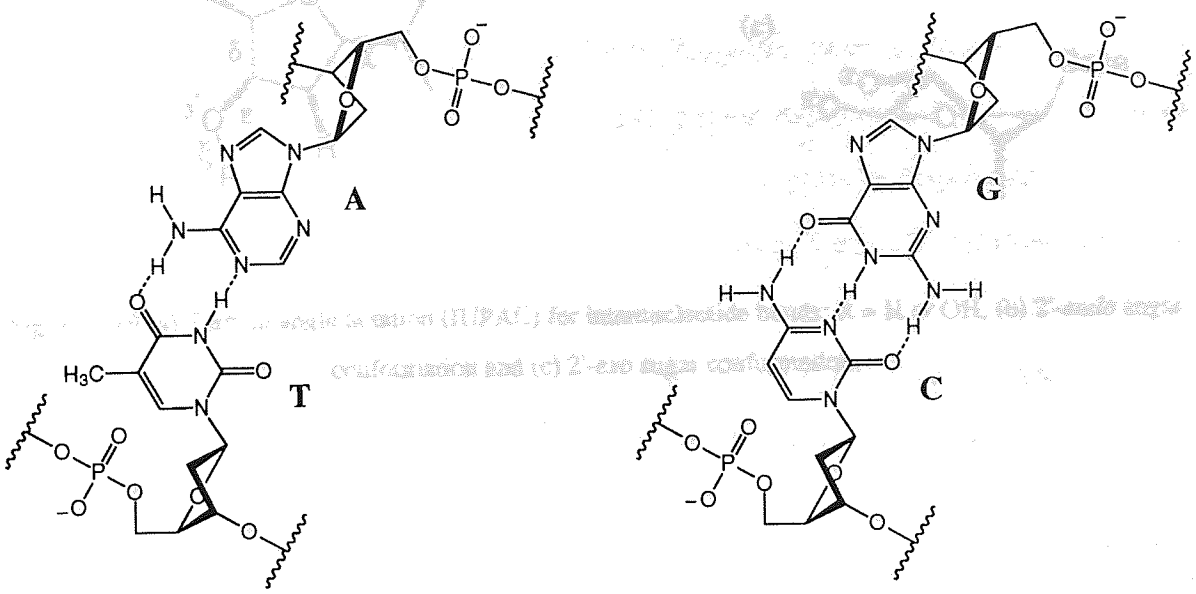


Figure 1.9 Hydrogen bonding between A·T and G·C base pairs.

1.6.1 Conformation of nucleotides

The repeating unit of the naturally occurring oligonucleotide polymer is the phosphodiester-linked nucleoside. The internucleotide bonds are assigned α, β, γ, δ, ε and ζ[86,87] (Figure 1.10 (a)). The conformation around the C4' and C5' bond determines the position of the 5'-phosphate relative to the sugar. The sugar ring puckers to minimise the non-bonding interactions between substituents and are assigned as C2'-*endo* or-*exo*. (Figure 1.10 (b), (c)). In solution these conformations are in equilibrium, separated by an energy barrier of less than 20 kJ mol^{-1}. The plane of the purine and pyrimidine heterocyclic bases are almost perpendicular to the sugar and approximately bisects the O4'-C1'-C2' angle occupying either the *syn* or *anti* conformations (Figure 1.11). Generally purines and pyrimidines occupy *anti*-

conformations, except guanine which is found in the *syn*-conformation as a mononucleotide; in alternating pyrimidine-purine sequence (eg CpGpCpGp) and in Z-DNA[88].

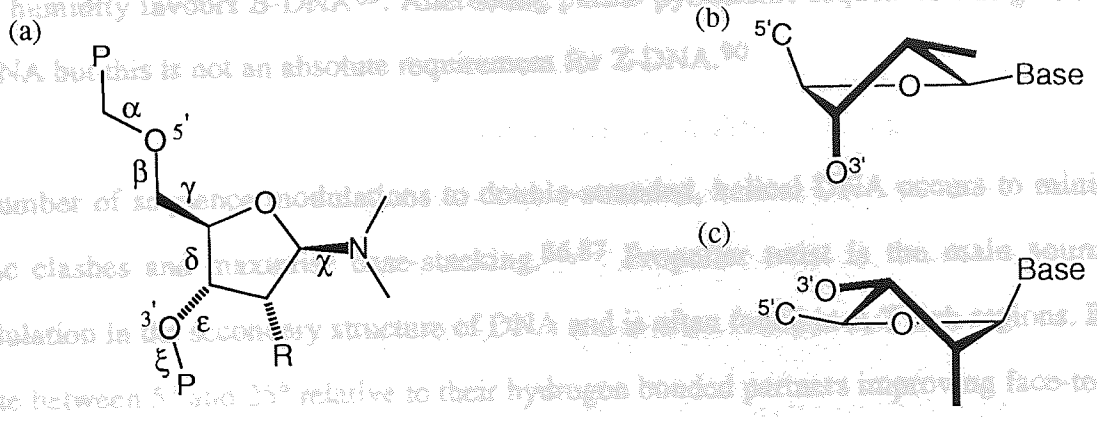


Figure 1.10 (a) Torsion angle notation (IUPAC) for internucleotide bonds; R = H or OH, (b) 2'-*endo* sugar conformation and (c) 2'-*exo* sugar conformation.

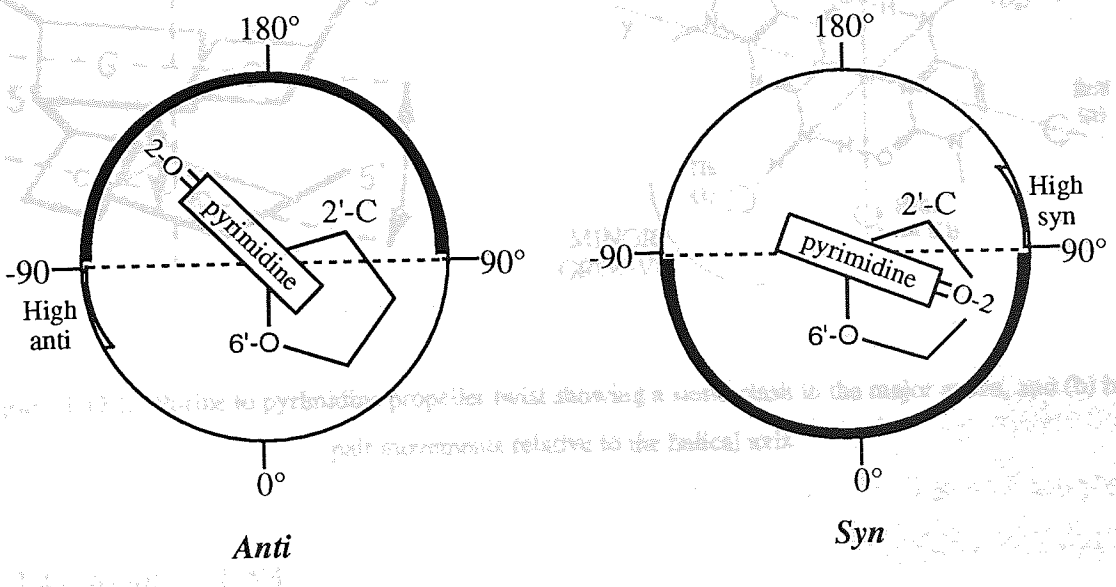


Figure 1.11 *Anti* and *syn* conformational ranges for pyrimidine nucleosides.

1.6.2 The secondary structure of DNA

The secondary structure of DNA is dependant on environmental conditions and on base sequence. At low humidity, crystal structures have revealed that A-DNA is preferred while high humidity favours B-DNA[89]. Alternating purine-pyrimidine sequences can give rise to Z-DNA but this is not an absolute requirement for Z-DNA.[90]

A number of sequence modulations to double-stranded, helical DNA occurs to minimise steric clashes and maximise base-stacking.[86,87] Propeller twist is the main source of modulation in the secondary structure of DNA and is often found in A·T rich regions. Bases rotate between 5° and 25° relative to their hydrogen bonded partners improving face-to-face contact between adjacent bases and, therefore, base-stacking. Steric clashes may result from propeller twists for pyrimidine-purine and purine-pyrimidine steps (Figure 1.12, (a)). To minimise the clashes caused by propeller twists, conformational changes in role (ρ), helical twist (Ω) tilt (τ), shift (D_x), slide (D_y) and rise (D_z) occur (Figure 1.12, (b)).

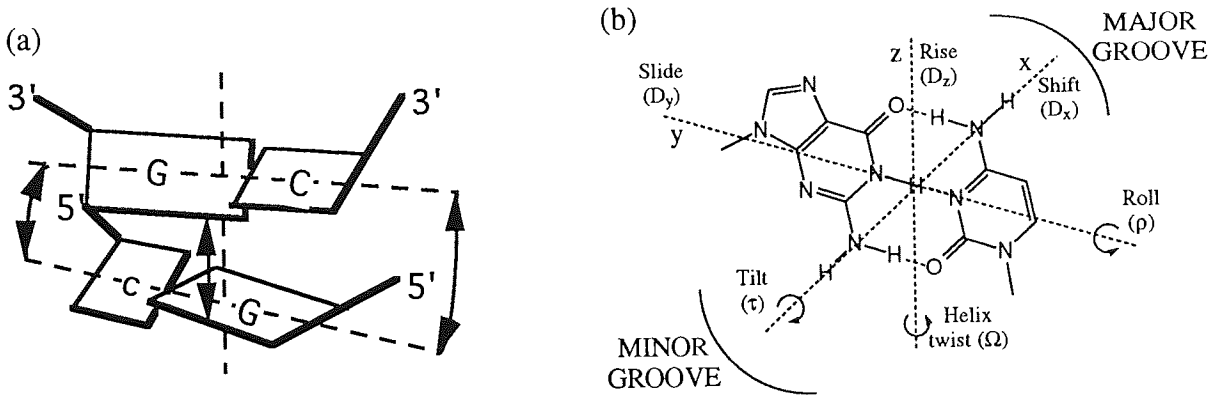


Figure 1.12 (a) Purine to pyrimidine propeller twist showing a steric clash in the major grove, and (b) base pair movements relative to the helical axis.

1.6.3 A-, B- and Z-DNA

Table 1.2 summarises the average helix parameters for the main DNA conformations. A-DNA exists in an antiparallel, right-handed (R) double helix.[87] The sugar rings are parallel to the helix axis, the phosphate backbone forms the circumference of the cylinder which is

240 nm in diameter. The sugar is C3' *endo*-puckered and the glycoside is in the *anti* conformation. The duplex has a hollow core along its axis 30 nm in diameter. The major groove is small but deep while the minor groove is wide but shallow. All DNA·RNA hybrids studied so far have always been in the A-conformation.

The major groove of B-DNA is comparable in depth to A-DNA but wider. Again, the sugar is C3' *endo*-puckered and the glycoside has the *anti* conformation. The major groove is covered with a monolayer of water, as is the phosphate backbone. The narrow, minor groove has a well ordered zig-zag chain of two water molecules per base pair.

	A-DNA	B-DNA	Z-DNA
Helix sense	R	R	L
Residues per turn	11	10	12
Twist per base pair/°	32.7	36	-9,-51
Displacement base pair / Å	4.5	-0.2,-1.8	-2-to-3
Rise per base pair / Å	2.56	3.3	3.7
Base tilt /°	20	-6	-7
Sugar pucker	3'-endo	2'-endo	3'-endo§
Minor groove width / Å	11.0	5.7	2.0
Major groove width / Å	2.7	11.7	8.8
Minor groove depth / Å	2.8	7.5	13.8
Major groove depth / Å	13.5	8.8	3.7

Table 1.2 Average parameters for A-, B- and Z-DNA. § Adopts the *syn*-conformation.

Z-DNA is left handed (L) with the duplex strands oriented antiparallel. One nucleoside in each base pair adopts the unusual *syn*-conformation, normally a purine. The *anti-syn* pattern alternates regularly along the DNA backbone. The sugar pucker alternates between C2'-*endo* and C3'-*endo* for each *syn*-oriented nucleoside. Other sub-types of DNA and RNA are of little biological importance, their secondary structure being the product of non-physiological environmental factors.

Crystal structures have revealed structured water associated with helical DNA. B-DNA in particular has been extensively studied[91,92]. The presence of structured water in the form of spines or ribbons in the minor groove of crystalline B-DNA appears to be related to groove width which is dependant on DNA sequence and crystalline environment. NMR studies have indicated that spines of hydration exist in solution. DNA hydration is thought to play an important role in DNA-ligand interactions. Hydration of duplex nucleic acids is further discussed in Sections 1.7.5 and 1.7.6.

1.7 Nature of Interactions Between Nucleic Acids

Hybridisation of two complementary strands of nucleic acid is a dynamic process which may be quantified by an equilibrium constant. Therefore, blockade of a biological nucleic acid target by an antisense oligonucleotide is never total. Section 2.1 discusses the rationale, advantages and disadvantages of modifying antisense oligonucleotides to improve their affinity for their target nucleic acid sequence.

1.7.1 Specificity and affinity of oligonucleotides for nucleic acid target sequences

The specificity of an oligonucleotide for its target sequence results from complementary hydrogen-bonding between bases of the antisense oligonucleotide and bases of the target nucleic acids. Base-pairing between sense and antisense nucleic acids generally follow the Watson-Crick hydrogen bonding motifs (Figure 1.9). The affinity of an oligonucleotide for its target sequence results from hybridisation interactions. The two major contributing factors to the free energy of hybridisation which stabilise nucleic acid duplexes are horizontal hydrogen bonds[66,86] (base sequence independent) and vertical stacking of the co-planar bases (base sequence dependant).[86,93] Affinity is a function of the base sequence and length of an oligonucleotide, a phenomenon exploited to calculate theoretical melting temperatures (T_m) for a given sequence by nearest neighbour methodology.[94,95] It has been shown that T_m depends linearly on the percentage content of GC base pairs if all other conditions, for example, pH, buffer media and salt concentration are constant.

1.7.2 Thermodynamic principles of duplex stability

Considered globally, the reversibility of duplex formation is the result of a negative but small free energy change (ΔG_{helix}) favouring helix formation (Equation 1).[96,97] Adding a base pair to a nucleic acid double helix leads to a small increment in its stability. The small negative free energy term per base pair is the result of opposing enthalpy and entropy changes. The global term ΔG_{helix} results from the relationship between enthalpy (ΔH_{helix}) and entropy (ΔS_{helix}) of hybridisation (Equation 4). The predominant factors stabilising duplex nucleic acids are hydrogen bonding (hb) and base-stacking interactions (s) although hydrophobic interactions (h) and conformational changes (r) are also thought to affect helix stability.

$$\Delta G_{helix} = \Delta G_r + \Delta G_h + \Delta G_s + \Delta G_{hb} \quad \text{[eqn. 1]}$$

$$\Delta H_{helix} = \Delta H_r + \Delta H_h + \Delta H_s + \Delta H_{hb} \quad \text{[eqn. 2]}$$

$$\Delta S_{helix} = \Delta S_r + \Delta S_h + \Delta S_s + \Delta S_{hb} \quad \text{[eqn. 3]}$$

Where $\Delta G_{helix} = \Delta H_{helix} - T\Delta S_{helix}$ [eqn. 4]

The major factors which influence enthalpic and entropic stabilisation of duplex nucleic acids are summarised in Figure 1.13. Each contribution is explained in more detail below.

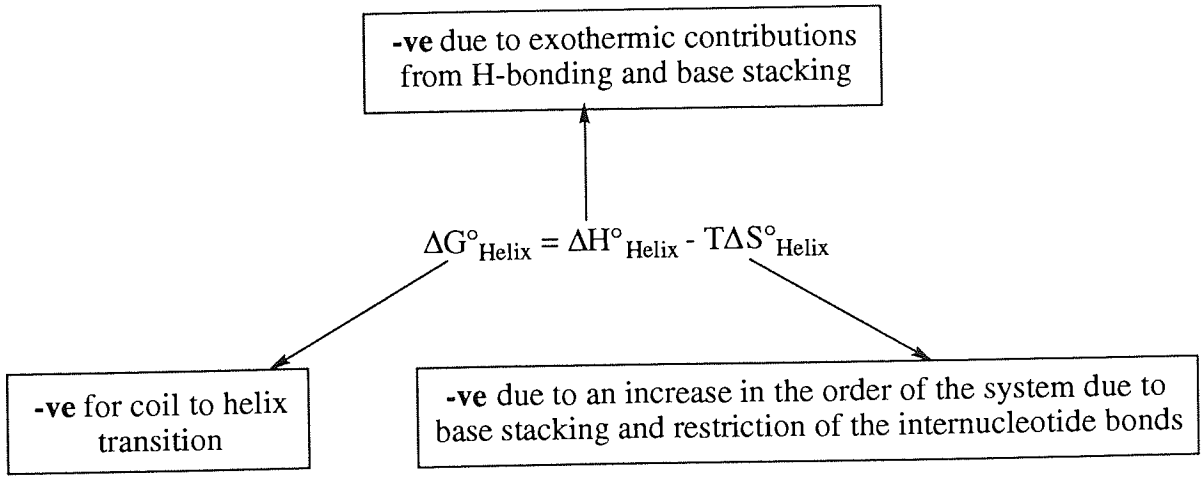


Figure 1.13 Summary of thermodynamic stabilising factors of duplex nucleic acids.

1.7.3 Base-stacking

Base-stacking in aqueous solution is exothermic and entropy consuming. Hunter et al 1990, suggested that base-stacking of aromatic systems occurs when the attractive interaction between π-electrons and the σ-framework outweigh unfavourable π-π repulsion[98]. Face-to-face stacked geometry is favoured by van der Waal's interactions and solvophobic effects but is generally disfavoured by π-π repulsion. The presence of polarising hetero atoms in, or attached to, an aromatic system has a major effect on electrostatic interactions, as is the case with purine and pyrimidine heterocycles. The geometry of the aromatic systems off-set to maximise π-σ interactions where unlike polarisations attract.

1.7.4 Hydrogen bonding

Hydrogen bonds are mainly electrostatic in character and cause predominantly enthalpic stabilisation of duplex nucleic acids independent of nucleoside sequence.[96,97] Hydrogen bonds have the general formula X-H···Y where X, Y are electronegative[85]. The strength of a hydrogen bond, reflected in its length, depends on charges located on X, H and Y. Hydrogen bonding between purine and pyrimidine heterocycles is of the type N-H···N and N-H···O with the donor N-H group being either the amino or imino type. Unlike covalent bonds which have well defined length, strength and orientation, hydrogen-bonds are 20 to 30 times weaker and are more susceptible to bending and stretching. Hydrogen-bonds are both additive and co-operative. Under the influence of a polarising hydrogen bond, the charge on X, H and Y adjust; H becomes more electropositive with X and Y become more electronegative. This effect leads to increased affinity of X, Y for accepting further hydrogen bonds.

1.7.5 Other interactions

Restriction of the six internucleotide bonds (Figure 1.10 (a)) incurs an entropic penalty. There is little experimental data regarding the thermodynamics of hydration specifically, largely due to the difficulties in separating effects due to hydration from those due to other causes. Hydrophobic effects are, however, thought to play little role in stabilising duplex

nucleic acids.[96,99] Optimal stacking of purine bases in a helix requires rolling, which induces propeller twist and thus steric clashes between bases. This repulsion can be relieved by geometrical adjustments e.g. rolling, tilting, helix twist angle and lateral sliding of the bases or base pairs (Figure 1.12). Such geometrical adjustments affects the enthalpy and entropy of base-stacking and the enthalpy of hydrogen bonding.

1.7.6 Environmental factors

Relative humidity, pH and counter cations influence the duplex stability of nucleic acids but are often kept constant during analysis of base-pairing and base-stacking. An aqueous medium is usually a prerequisite for duplex formation of nucleic acids since organic solvents with low dielectric constant denature nucleic acid duplexes. DNA is relatively stable between pH 5 and 9.[86,96]

The relative humidity and ionic strength of the aqueous medium determines the polymorphic character of DNA duplexes. Unlike proteins which have a hydrophobic core, nucleic acids do not have a clearly defined interior. DNA duplexes have a tightly bound hydration shell, having on average about 20 water molecules per nucleoside which is not very permeable to ions.[100] The secondary hydration shell is poorly defined and has some characteristics of bulk water. High relative humidity favours the B conformation and reduced relative humidity (or increased ionic strength) renders B to A or B to Z conformational transitions.

The affect of counter ions on nucleic acid stability is not clear although T_m rises with salt concentration. With the exception of lithium, no specific interaction between monovalent cations and DNA has been found.[101] This is also true for divalent organic cations such as polyamine and methonium ions. The effects of divalent ions such as calcium and magnesium on nucleic acid duplex stability are more complex and not fully understood. Observations suggest that the enthalpy difference between coil and helical states is

independent of sodium ion concentration, suggesting that the increase in T_m with ion concentration is entropic in origin.[102]

1.8 Mechanism of Action of Antisense Oligonucleotides

The majority of traditional biological targets for medicinal chemists have been proteins, either receptors or enzymes. Antisense oligonucleotides can be used to modify cellular function by controlling protein production at the nucleic acid level of transcription and translation (Figure 1.14).[67,103-105]

1.8.1 Inhibition of transcription

Transcription of genetic information from double-stranded DNA to single-stranded mRNA starts with RNA polymerase binding to a promoter (start sequence) on helical DNA. One turn of helical DNA is unwound where one strand of DNA, determined by the promoter sequence, acts as a template for RNA polymerase to synthesise mRNA. RNA polymerase moves along the DNA helix extending the mRNA molecule until the termination signal is reached. The 5'-end of the RNA molecule, where the terminal sugar is 2'-methylated, is capped with a charged 7-methylguanosine connected *via* a 5' to 5'-triphosphate bridge (+Gppp).[105] Antisense oligonucleotides can hybridise to the unwound DNA and so act to block RNA polymerase sterically. Antisense oligonucleotides directed against the TAR region of the HIV genome can potentially inhibit tat-TAR interactions, blocking or slowing viral replication.[103]

1.8.2 Inhibition of post-transcriptional processes

The post-transcriptional transcript, pre-mRNA, is subject to a number of maturation processes before translocation to the cytoplasm for translation. Non-coding introns are spliced out, the 5'-terminus is capped and various bases are altered. The 3'-terminus is polyadenylated which is associated with the export of mRNA into the cytoplasm. Formation of double-stranded RNA, especially in the region of the polyadenylation site, prevents translocation of mRNA into the cytoplasm.[72]

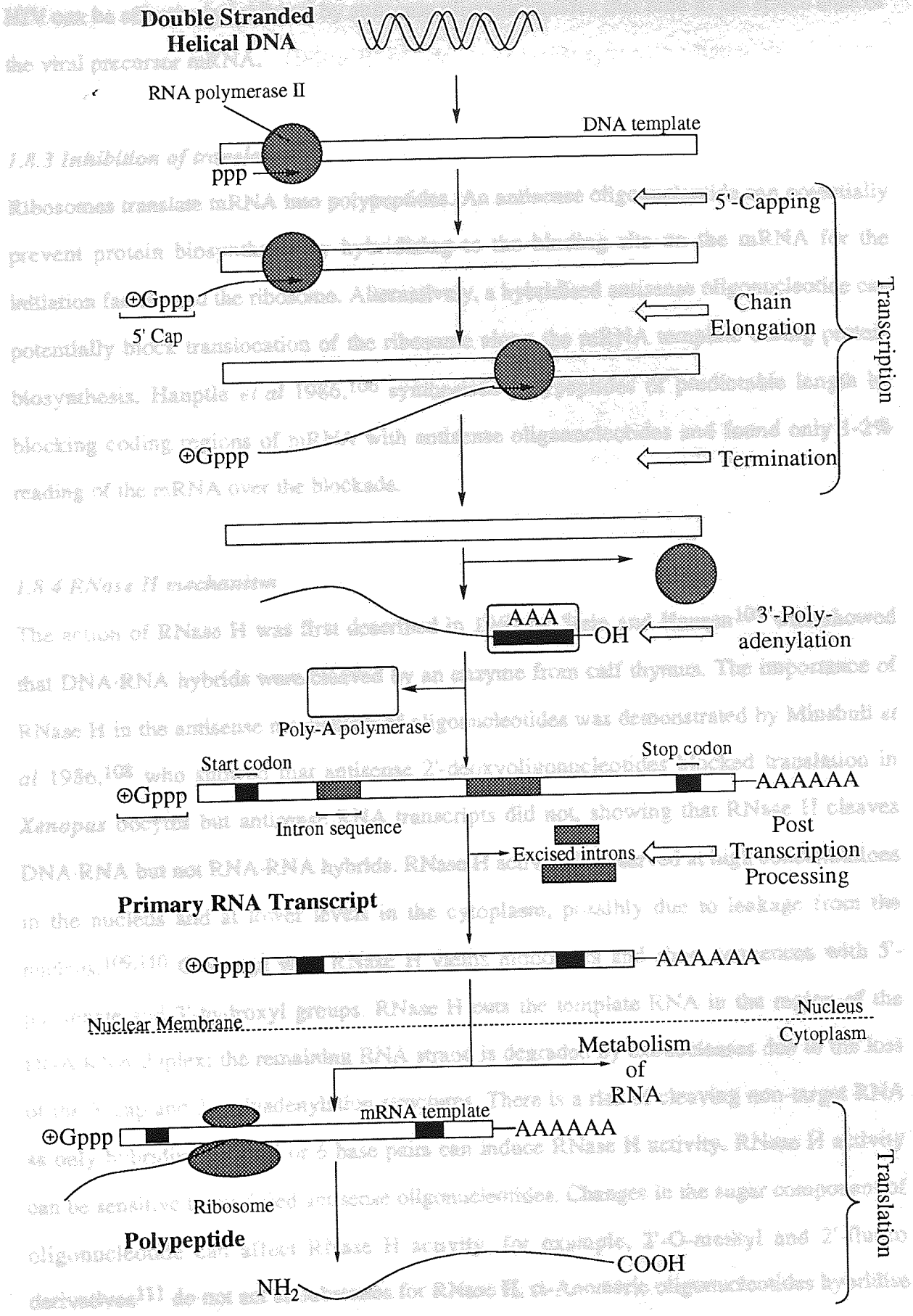


Figure 1.14 RNA processing.

HIV can be effectively inhibited by antisense oligonucleotides that bind to the splice sites of the viral precursor mRNA.

1.8.3 Inhibition of translation

Ribosomes translate mRNA into polypeptides. An antisense oligonucleotide can potentially prevent protein biosynthesis by hybridising to the binding site on the mRNA for the initiation factors and the ribosome. Alternatively, a hybridised antisense oligonucleotide can potentially block translocation of the ribosome along the mRNA template during protein biosynthesis. Hauptle *et al* 1986,[106] synthesised polypeptides of predictable length by blocking coding regions of mRNA with antisense oligonucleotides and found only 1-2% reading of the mRNA over the blockade.

1.8.4 RNase H mechanism

The action of RNase H was first described in 1969 by Stein and Hausen[107] who showed that DNA·RNA hybrids were cleaved by an enzyme from calf thymus. The importance of RNase H in the antisense mechanism of oligonucleotides was demonstrated by Minshull *et al* 1986,[108] who showed that antisense 2'-deoxyoligonucleotides blocked translation in *Xenopus* oocytes but antisense RNA transcripts did not, showing that RNase H cleaves DNA·RNA but not RNA·RNA hybrids. RNase H activity is observed at high concentrations in the nucleus and at lower levels in the cytoplasm, possibly due to leakage from the nucleus.[109,110] Cleavage with RNase H yields monomers and short sequences with 5'-phosphate and 3'-hydroxyl groups. RNase H cuts the template RNA in the region of the DNA·RNA duplex; the remaining RNA strand is degraded by exonucleases due to the loss of the 5'-cap and 3'-polyadenylation structures. There is a risk of cleaving non-target RNA as only hybridisation of 5 or 6 base pairs can induce RNase H activity. RNase H activity can be sensitive to modified antisense oligonucleotides. Changes in the sugar component of oligonucleotide can affect RNase H activity, for example, 2'-O-methyl and 2'-fluoro derivatives[111] do not act as substrates for RNase H. α-Anomeric oligonucleotides hybridise to RNA but do not induce RNase H activity[112]. Modifications to the oligonucleotide

backbone can also influence RNase H activity. Phosphorothioate oligonucleotides, for example, are substrates for RNase H[113] while methylphosphonate oligonucleotides are not.[114]

Chapter 2

Design and Synthesis of Modified Nucleosides for Antisense Applications

2.1 Rationale

Hybridisation of an antisense oligonucleotide with its nucleic acid target, in aqueous solution, is a dynamic process which reaches a point of equilibrium for a particular set of environmental conditions (Figure 2.1). By their very nature, at any given moment, antisense oligonucleotides cannot deliver total blockade of a nucleic acid target sequence.

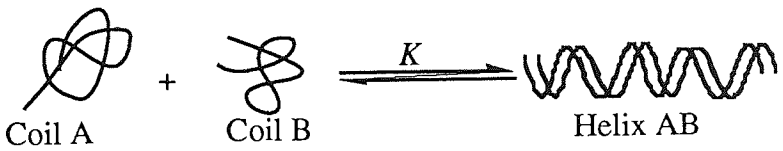


Figure 2.1 Coil-to-helix transition.

The balance between the concentrations of the unhybridised coils and hybridised helical nucleic acid is defined by the equilibrium constant *K*. (Equation 5). The system will adjust to compensate for varying concentrations of the antisense coil, the target coil and the antisense·target hybrid duplex so that *K* remains constant.

$$K = \frac{[\text{Helix A} \cdot \text{B}]}{[\text{Coil A}][\text{Coil B}]} \qquad \text{[Eqn. 5]}$$

Increasing the affinity of an antisense oligonucleotide for its target nucleic acid sequence, by structurally modifying the antisense oligonucleotide, will make the resulting helical hybrid more thermodynamically stable. A new equilibrium constant will thus be defined since the position of the equilibrium shifts towards the helical state. *In vitro*, the modified antisense oligonucleotide with increased affinity for its target nucleic acid sequence will

achieve a better blockade at equivalent intracellular concentration, compared to the unmodified antisense oligonucleotide. If the susceptibility of the helical hybrid of the modified antisense oligonucleotide and target nucleic acid towards helical melting by cellular processes is lowered, it follows that a lower intracellular concentration of the modified antisense oligonucleotide would be required to achieve an effective blockade of the nucleic acid sequence. This is potentially advantageous as cellular uptake of antisense oligonucleotides is an inefficient process and intracellular concentrations are low.

The potential risk associated with this 'high affinity' strategy is the loss of specificity of the modified antisense oligonucleotide for its target nucleic acid sequence due to potential base-pair mismatches. Nucleic acid duplexes are destabilised by base-pair mismatches but the duplex can be potentially stabilised thermodynamically by the improved, sequence-independent affinity of the modified antisense oligonucleotide for a non-targeted nucleic acid sequence resulting from base-stacking interactions.

2.2 Literature Precedents

2.2.1 Aralkylation of hexose DNA is thermodynamically stabilising

Entropic stabilisation of hexose-DNA duplexes can be achieved by incorporating the aralkylated nucleoside, 2-benzyloxyadenosine (A.BnO), into hexose oligonucleotides.[115] In hexose DNA, isoguanine (I) forms a Watson-Crick purine-purine base pair with guanine (G). The self-complementary sequence I_3G_3, associates in aqueous solution forming a duplex with a melting temperature T_m = 57°C. Benzylation at O2 of isoguanine giving A.BnO which has a marked effect on duplex stability. The comparable self-complementary sequence $A.BnO_3G_3$ has a melting temperature T_m = 71°C.[116]

It is postulated that increased thermodynamic stability of the self-complementary sequence $A.BnO_3G_3$ is due to enhanced base-stacking interactions associated with the conjugated π systems of the aralkyl substituents and purine heterocycles. The hydrophobicity of the

aralkyl substituents may contribute to greater thermodynamic duplex stability through entropic changes.

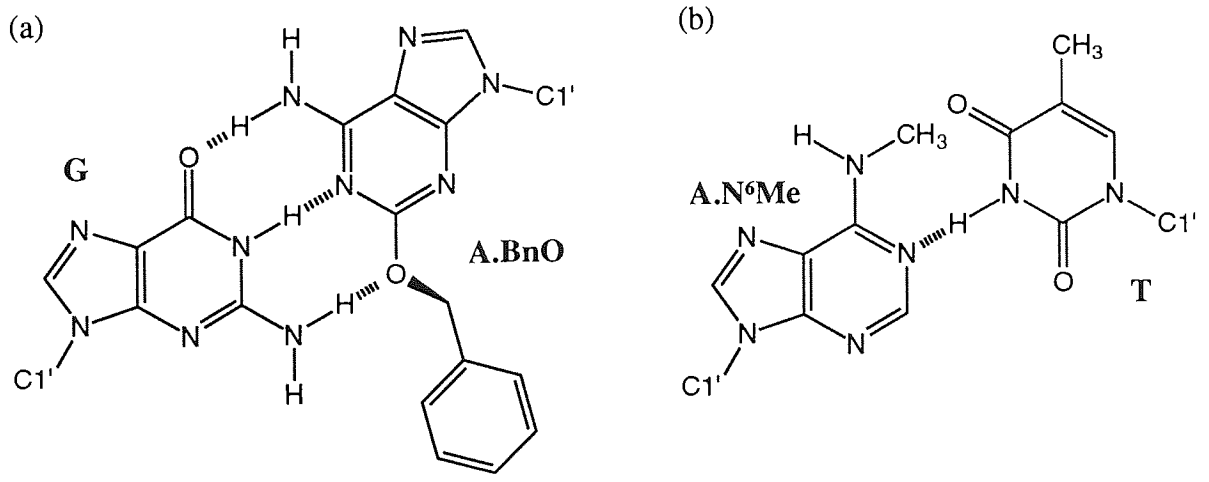


Figure 2.2 Possible conformations of the (a) hexose 2-benzyloxyadenosine-guanine base-pair and (b) pentose 6-methyladenosine-thymine base pairs.

The literature shows that inclusion of N6-methylated adenosine in pentose-DNA sequences results in duplex destabilisation.[117,118] N6-Methylation of adenosine (A.N^{6}Me)has an important biological role in gene regulation. Restricted rotation around the C6-N6 bond of N6-methylated adenosine is caused by conjugation of the N6 amine lone pair of electrons with the purine π-system (Figure 2.2 (b)). The methyl group of A.N^{6}Me is preferentially found in the *cis* orientation with respect to N1 and so is unable to form a competent hydrogen bond with O4 of thymine.

2.2.2 Can pentose DNA be thermodynamically stabilised by aralkyl substituents?

These observations from aralkyl-substituted hexose-DNA warranted an investigation of pentose-DNA containing aralkyl substituents. An N4-aralkylated analogue of cytosine, namely: 1-(2'-Deoxy-β-D-*erythro*-pentofuranosyl)-4*N*-phenethyl-pyrimidine-2-one (**23**), was initially considered as a representative example. Steric restrictions within the oligonucleotide duplex containing **23** (Figure 2.3 (a), denoted as **C.N^{4}EtPh**) would be expected to force the bulky aralkyl groups away from the hydrogen bond acceptor of

guanine, overcoming any rotational restrictions about the C4-N4 bond of the purine heterocycle, possibly incurring some energetic penalty. This would potentially allow the modified cytosine **C.N^{4}EtPh** to form a Watson-Crick H-bond base-pair with guanine. The *+I* effect of N4-aralkylation would be expected to weaken H-bonding between the NH of the modified cytosine **23** and the O6 of guanine compared with unmodified cytosine. UV spectrophotometry shows that N6-methylation of adenosine (A.N^{6}Me) causes a bathochromic shift of 10 nm, indicating that the electronic nature of the nucleoside has been altered compared to adenosine. Consequently, this could affect the base-stacking potential of A.N^{6}Me with bases above and below its plane within oligonucleotides. To assess the effect of aralkyl substituents in the major groove of pentose-DNA in terms of thermal stability, modified cytosines, for example **C.N^{4}EtPh** (Figure 2.3 (a)), were incorporated into short oligonucleotides using conventional, automated DNA solid phase synthesis. The presence of an NH group at C4 of the modified cytosine **C.N^{4}EtPh** ensures that a Watson-Crick hydrogen bonding motif with thymine is still achievable given appropriate steric and energetic conditions.

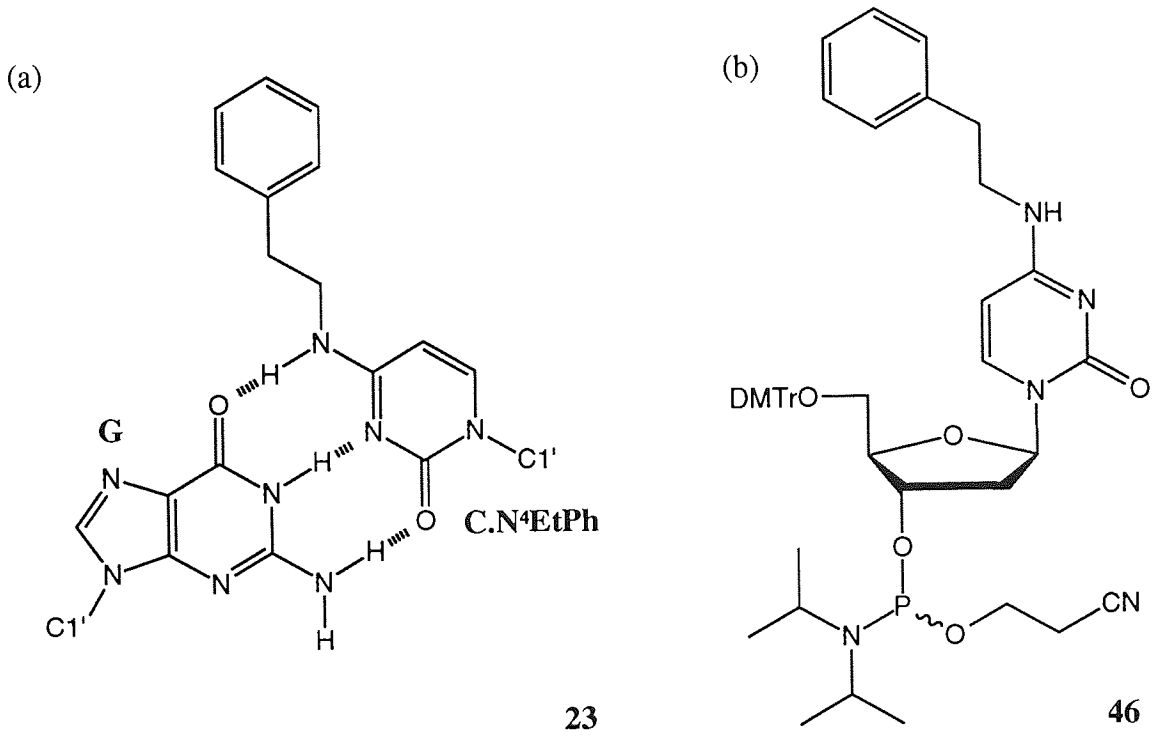


Figure 2.3 (a) Possible Watson-Crick H-bond motif between C4-aralkyl-cytosine and guanine and (b) protected cytosine nucleoside **46** for automated solid phase synthesis.

DNA solid phase synthesis requires 4,4'-dimethoxytritylation of the parent nucleoside 5'-hydroxyl and phosphitylation of the 3'-hydroxyl to form a 3'-[(2-cyanoethyl)-*N*,*N*-di*iso*propyl]phosphoramidite, such as **46**. All other reactive amine and alcohol functions require protection with acid-stable, base-labile groups.[65]

2.3 Literature review of pyrimidine C4 aralkylation strategies

In order to introduce aralkyl groups at N4, a pyrimidine nucleoside intermediate with a displaceable group a C4 was sought. The C4 adduct of the intermediate nucleoside should be susceptible to displacement at C4 by primary amine nucleophiles, using relatively mild conditions, to yield N4 aralkyl-substituted cytosines.

The literature showed that several strategies have been employed to make C4-aralkyl pyrimidine nucleosides and C6-aralkyl purine nucleosides. The common starting materials for these reactions are usually uridine, 2'-deoxyuridine, guanine and 2'-deoxyguanosine, following the general procedure:

i. formal or transient protection of sugar hydroxyls and exocyclic amine functions;
ii. activation of the pyrimidine O4 or purine O6 amide carbonyls;
iii. *in situ* formation of N4 pyrimidine or N6 purine quaternary nitrogen salts;
iv. attack on the quaternary nitrogen salt by a nucleophile to form stable C4 or C6-substituted intermediates which can be isolated;
v. subsequent formation of C4-aralkylated pyrimidine nucleosides or C6-aralkylated purine nucleosides using amine or hydroxyl nucleophiles;
vi. deprotection of hydroxyl and amine functions and isolation of the final product.

Sung (1981)[119] described the synthesis of 1,2,4-triazolyl-thymine **2** which was shown to be a useful intermediate in the synthesis of 5-methylcytosine **3** (Figure 2.3). Acetyl protecting groups were used to mask the reactive 5'- and 3'-hydroxyls of 2'-deoxythymine to prevent unwanted side reactions **1**. Activation of the 4-carbonyl of **1** was achieved using 4-

chlorophenyl phosphorodichloridate (1.5 eq.) in pyridine. 1,2,4-Triazole (3 eq.) was added to the same reaction pot and the reaction mixture stirred for three days to give the C4-substituted azole **2** in 78% yield. Treatment of **2** with dioxane-ammonia (3:1 $^{v}/_{v}$) facilitated displacement of the C4-triazolyl group by ammonia and concomitant deprotection of the sugar alcohol functions yielded 5-methyl cytosine **3** in 58% yield.

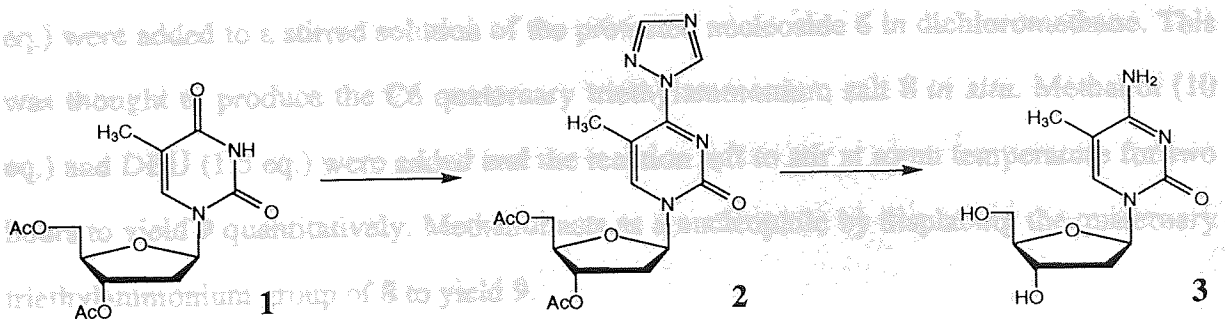


Figure 2.4 Synthesis of 5-methyl thymine **3**.

In a paper published in the same year, Sung (1981)[120] described the synthesis of a 3'-monophosphate-protected thymine nucleoside **5** suitable for bench synthesis of oligonucleotides via the phosphotriester method (Figure 2.5). The 5'-OH of 2'-deoxyuridine was protected with 4,4'-dimethoxytrityl chloride to give **4**. 4-Chlorophenyl phosphorodichloridate was used to activate the C4-carbonyl of the pyrimidine heterocycle and to phosphorylate the 3'-OH of the sugar of **4**. After addition of 1,2,4-triazole, the reaction was stirred for four days after which 2-cyanoethanol was added giving the 3'-phosphotriester **5** in 65% yield.

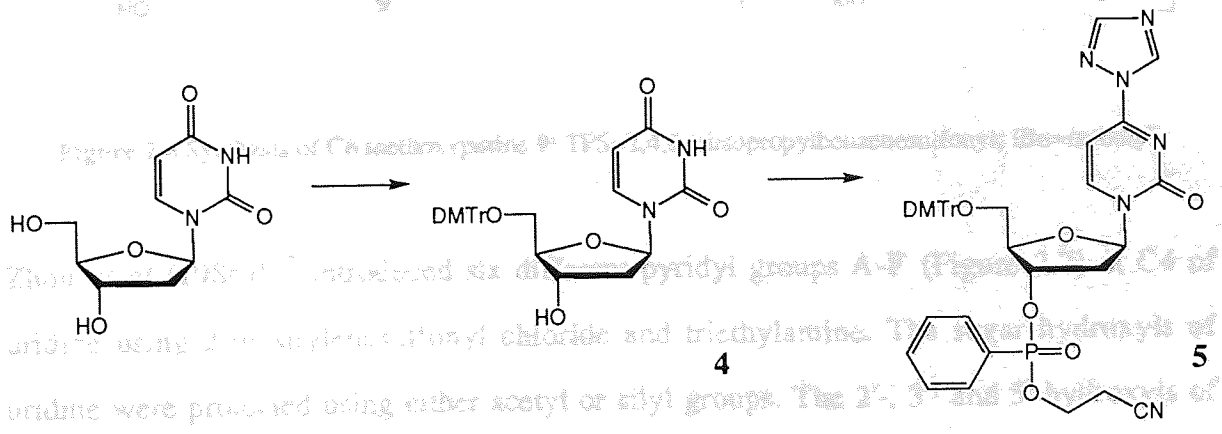


Figure 2.5 Synthesis of **5**, suitable for synthesis of oligonucleotides.

Gaffney and Jones (1982)[121] synthesised the sugar-protected O6-sulfonyl derivative **7** of 2'-deoxyguanosine to study the susceptibility of C6 to attack by alcohol nucleophiles. Addition of methanol resulted mainly in the formation of the parent 3',5'-*O*-di*iso*butyryl-nucleoside starting material **6** due attack on sulphur. It was known that sulfonate functions underwent displacement with amine nucleophiles. Trimethylamine (4 eq.) and DMAP (0.05 eq.) were added to a stirred solution of the protected nucleoside **6** in dichloromethane. This was thought to produce the C6 quaternary triethylammonium salt **8** *in situ*. Methanol (10 eq.) and DBU (1.5 eq.) were added and the reaction left to stir at room temperature for two hours to yield **9** quantitatively. Methanol acts as a nucleophile by displacing the quaternary triethylammonium group of **8** to yield **9**.

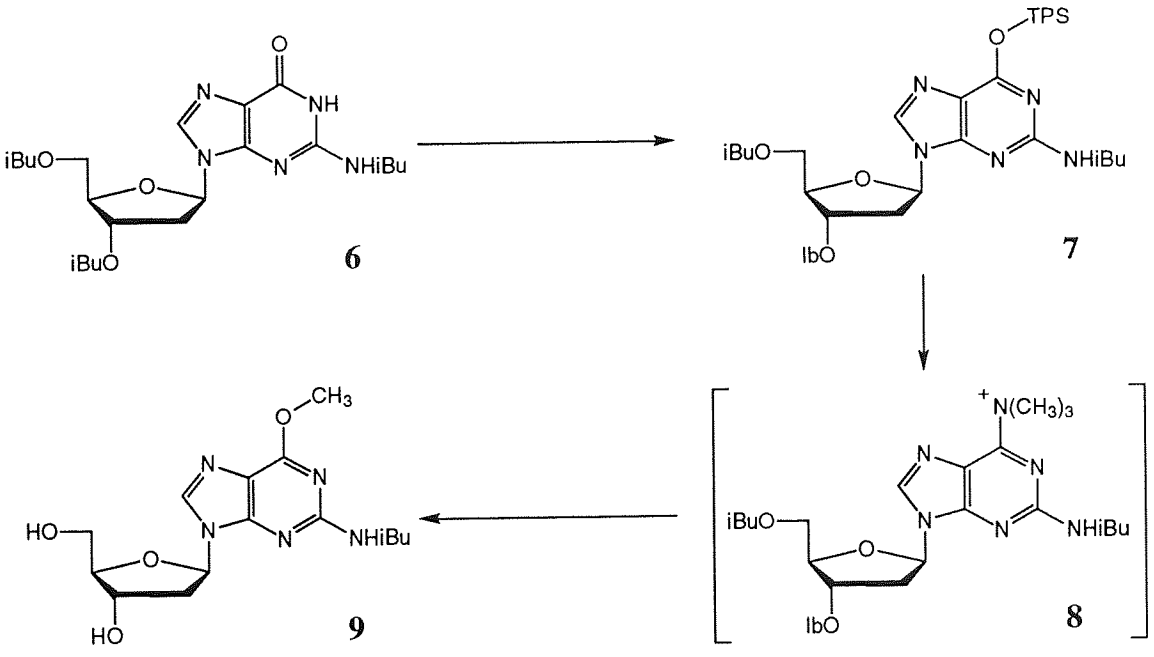


Figure 2.6 Synthesis of C6 methoxypurine **9**; TPS=2,4,6-tri*iso*propylbenzenesulfonyl; iBu=*iso*butyl.

Zhou *et al* (1986)[122] introduced six different pyridyl groups **A-F** (Figure 2.7) at C4 of uridine using 2-mesitylenesulfonyl chloride and triethylamine. The sugar hydroxyls of uridine were protected using either acetyl or silyl groups. The 2'-, 3'- and 5'-hydroxyls of uridine were protected as acetyl esters giving **10**. The conditions necessary to remove the acetyl protecting groups of nucleosides **11A-11F** are incompatible with the base-sensitive

nature of the pyridyl groups at C4. Fluoride-labile, silyl protection of the sugar hydroxyls of uridine was successfully employed, thus avoiding the need for aqueous basic deprotection conditions. The 3'- and 5'-hydroxyls of uridine were protected with the cyclic 3',5'-*O*-(1,1,3,3-tetra*iso*propylsiloxan-1,3-diyl) group (**12**) and the 2'-hydroxyl transiently protected using trimethylsilyl chloride. Activation of the C4 carbonyl of **10** and **12** was achieved using 2-mesitylenesulfonyl chloride (3 eq.) and triethylamine (10 eq.) in dichloromethane, forming a C4 triethylammonium salt which was not isolated. Addition of pyridyl heterocycles **A-F** to the C4 triethylammonium salt of **10** yielded **11A-11F**, whereas addition of the pyridyl heterocycles **A-F** to the C4 triethylammonium salt of **11**, followed by desilylation using TBAF, yielded **13A-13F**.

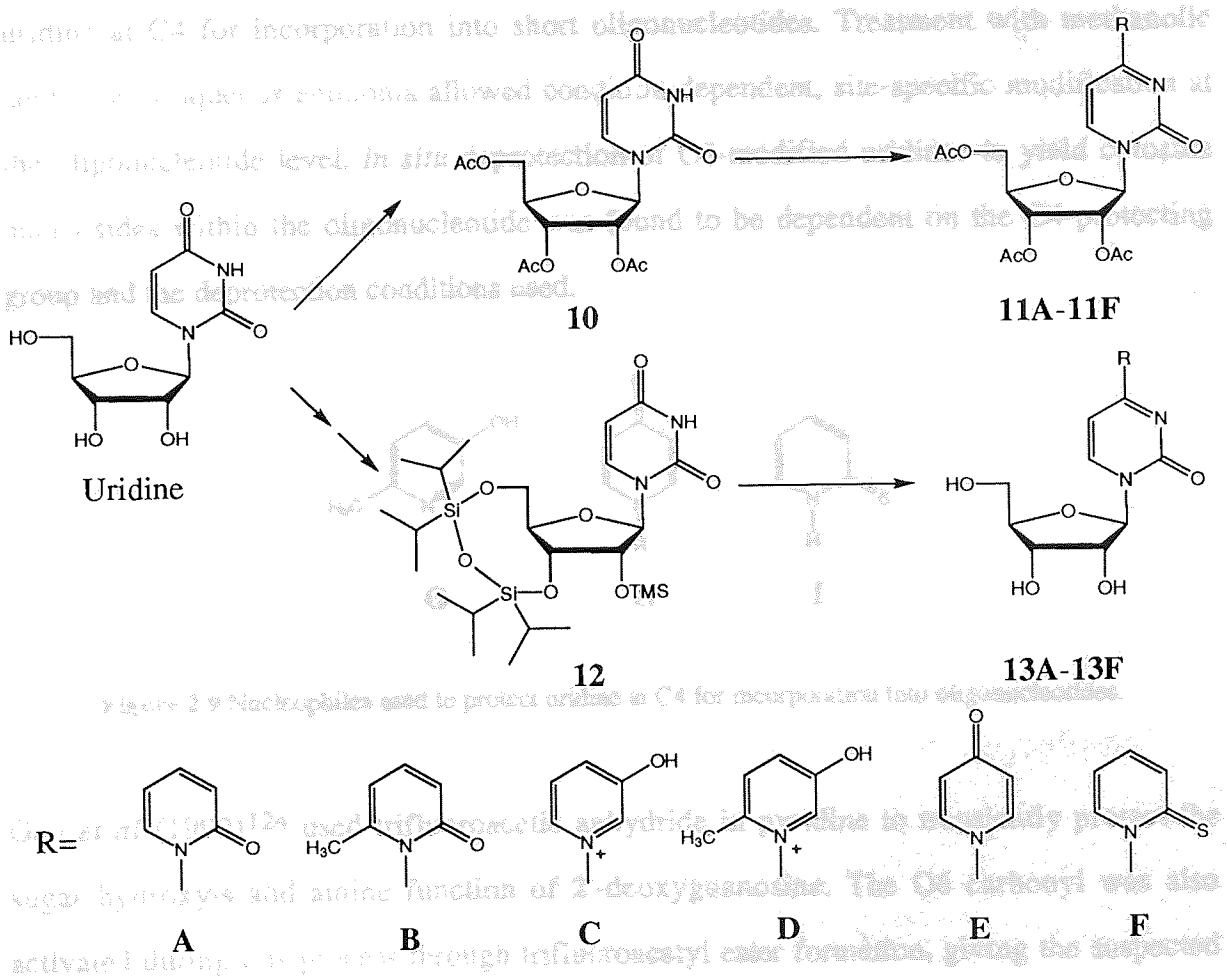


Figure 2.7 Synthesis of C4 pyridine derivatives **11A-11F** and **13A-13F**; TMS = trimethylsilyl.

Nucleosides **11A-11F** and **13A-13F** provided stable, C4 pyridyl-substituted pyrimidine nucleosides, which were all shown to undergo attack by amine nucleophiles (Figure 2.8) in good yield.

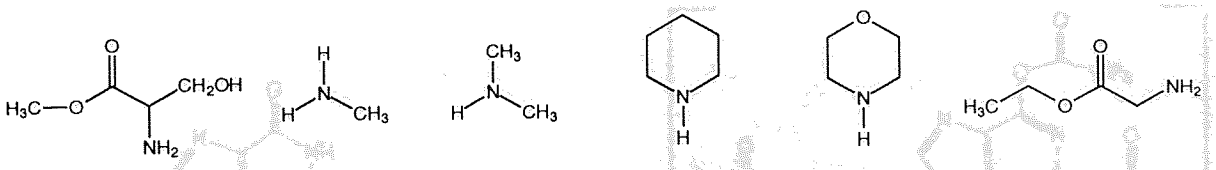


Figure 2.8 Nucleophiles used to displace the C4 pyridyl function of nucleosides **11A-11F** and **13A-13F**.

Further work by Zhou *et al* (1986),[123] used nucleophiles **G**, **H**, and **I** (Figure 2.8) to protect uridine at C4 for incorporation into short oligonucleotides. Treatment with methanolic ammonia or aqueous ammonia allowed condition-dependent, site-specific modification at the oligonucleotide level. *In situ* deprotection of C4-modified uridines to yield cytosine nucleosides within the oligonucleotide was found to be dependent on the C4-protecting group and the deprotection conditions used.

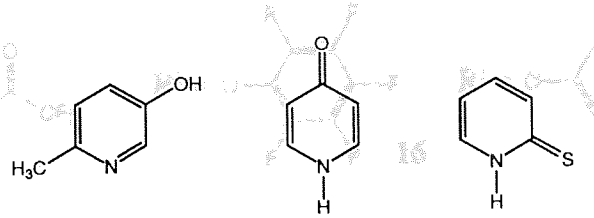


Figure 2.9 Nucleophiles used to protect uridine at C4 for incorporation into oligonucleotides.

Gao *et al* (1992)[124] used trifluoroacetic anhydride in pyridine to transiently protect the sugar hydroxyls and amine function of 2'-deoxyguanosine. The O6 carbonyl was also activated during this process through trifluoroacetyl ester formation, giving the suspected intermediate **14** (Figure 2.10). The C6 centre was thought to undergo nucleophilic attack by pyridine to form a N6 pyridyl intermediate **15**, which was not isolated. Addition of 4-nitrophenol or pentafluorophenol to the same reaction pot yielded **16** or **17** with

concomitant deprotection of the 3'- and 5'-hydroxyl groups in 95% and 67% yields respectively. Nucleosides **16** and **17** were shown to undergo attack by a range of amine nucleophiles.

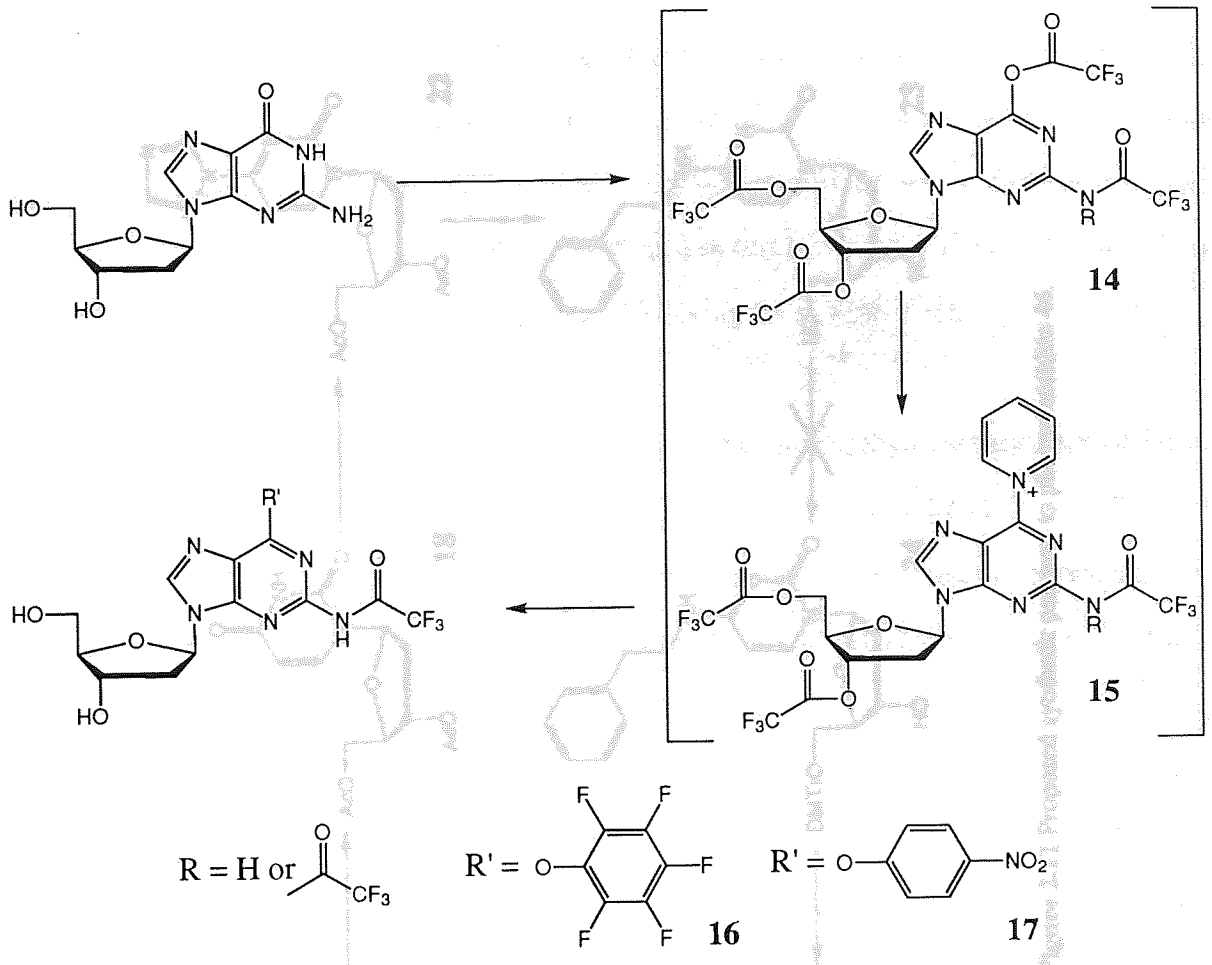


Figure 2.10 Synthesis of C6-protected purine nucleosides **16** and **17**.

2.4 Synthesis of Modified Pyrimidines for Automated DNA Synthesis

2.4.1 Sugar protection and C4 aralkylation of uridine

Initial attempts to synthesise **46** were based on the synthesis of **2** as a means of introducing the phenethyl group at C4 (Figure 2.11). Synthesis of **22** was achieved following the method of Sung[119], using 4-Chlorophenyl phosphorodichloridate to activate C4 of 2'-deoxyuridine.

Figure 2.11 Proposed synthetic pathway to phosporamidite **46**.

To prevent phosphorylation of the 3'- and 5'-hydroxyls, formal protection of these functions was first required. Many protecting agents have been employed to mask sugar hydroxyls. These include acetyl, benzoyl, di-*tert*-butylmethylsilyl, trityl and 4,4'-dimethoxytrityl (DMTr) groups.[119-122,125-127] Acetyl esters were used to protect the 3'- and 5'-hydroxyl functions due the ease of their removal using saturated methanolic ammonia. In this work, 2'-deoxyuridine was co-evaporated three times in pyridine, placed under argon, dissolved in pyridine and cooled to 0°C. To this, a catalytic amount of DMAP was added followed by acetic anhydride (5 eq.) and the resulting solution was left to stir at 4°C for 12 hours. The product was isolated by flash chromatography,[128] giving **18** in 81% yield.

4-Chlorophenyl phosphorodichloridate (2 eq.) was added to an anhydrous solution of **18** in pyridine at 0°C. After 5 minutes, 1,2,4-triazole (4 eq.) was added and the mixture left to stir at room temperature for 72 hours. Purification was achieved by flash chromatography, followed by crystallisation from ethyl acetate, giving **22** as colourless crystals in 53% yield.

The mechanism of the formation of intermediate quaternary ammonium salts of C4 pyrimidine nucleosides had been speculated upon in early publications. Rysard *et al* (1986)[129] followed the reaction of 2',3',5'-tri-*O*-triacetylinosine with 4-Chlorophenyl phosphorodichloridate by ^{31}P-NMR and isolated the resulting N6 quaternary pyridinium salt. ^{31}P-NMR showed activation of the C6 carbonyl by 4-Chlorophenyl phosphorodichloridate occurred. The proposed mechanism for the synthesis of C4 triazolyl pyrimidine nucleoside **22** is illustrated in Figure 2.12.

The C4 azole of **22** proved to be readily displaced by the amine nucleophile phenethylamine (1 eq.) in 1,4 dioxane at 70°C overnight. The product solution was concentrated under vacuum and the residue dissolved in saturated methanolic ammonia to deprotect the 3'- and 5'-acetyl-protected hydroxyl functions. Nucleoside **23** was isolated as a foam after purification by flash chromatography in 85% yield (Figure 2.10).

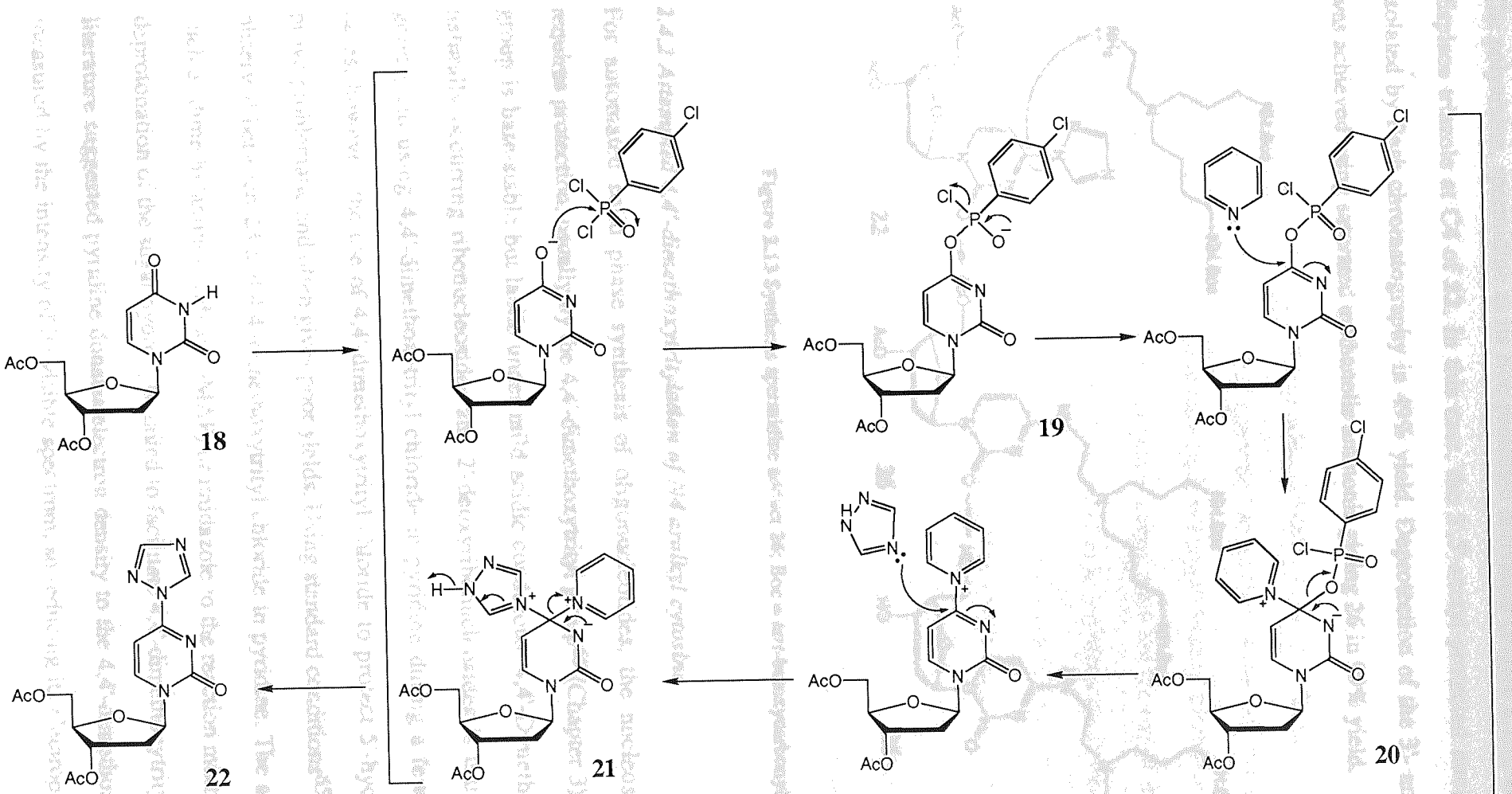


Figure 2.12 Proposed mechanism for the synthesis of C4 triazolyl pyrimidine nucleoside **22** from **18**

The polyamine, spermidine as its Boc-protected derivative (Figure 2.13), was also used to displace triazole at C4 of **22**. In this case, the 3',5'-acetyl-protected nucleoside **25** was isolated by flash chromatography in 49% yield. Deprotection of the 3'- and 5'-hydroxyls was achieved using saturated methanolic ammonia giving **26** in 99% yield.

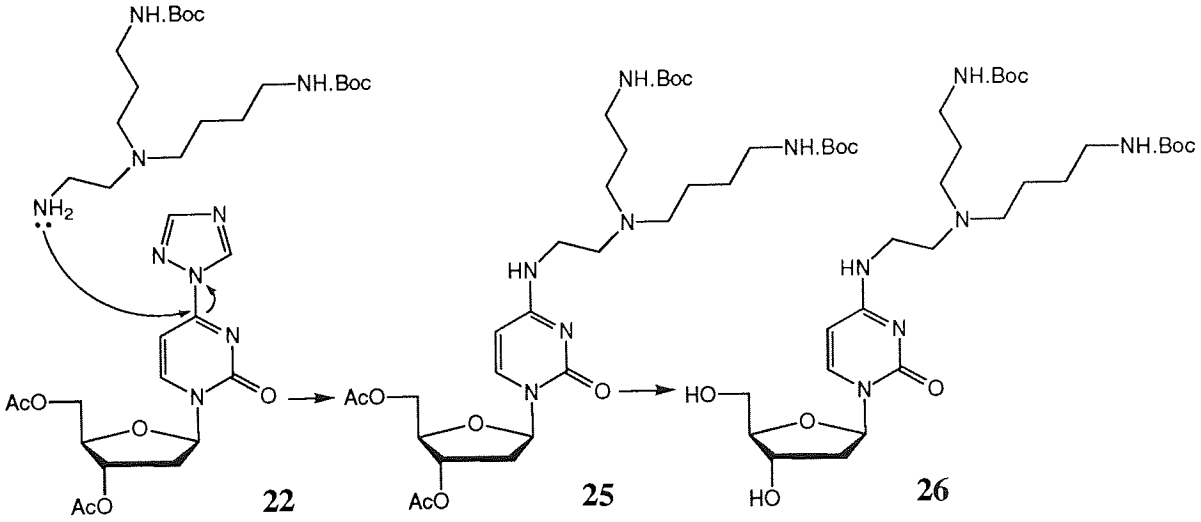


Figure 2.13 Synthesis spermidine adduct **26**; Boc = *tert*-butoxycarbonyl.

2.4.2 Attempted 4,4'-dimethoxytritylation of N4 aralkyl cytosine

For automated solid phase synthesis of oligonucleotides, the nucleoside 5'-hydroxyl requires protection, usually by the 4,4'-dimethoxytrityl group[65] (Chapter 3). This protecting group is base-stable but labile under mild acidic conditions. 4,4'-Dimethoxytritylation of naturally occurring ribonucleosides and 2'-deoxyribonucleosides is usually achieved in good yield using 4,4'-dimethoxytrityl chloride in pyridine during a few hours. In some cases, however, the use of 4,4'-dimethoxytrityl chloride to protect 5'-hydroxyl group can prove problematic and often gives poor yields. Using standard conditions[65], no reaction was observed between **23** and 4,4'-dimethoxytrityl chloride in pyridine. The addition of a base such as dimethylaminopyridine (DMAP) or imidazole to the reaction mixture, to encourage deprotonation of the sugar hydroxyls, failed to facilitate 4,4'-dimethoxytritylation of **23**. The literature suggested pyridine donates electron density to the 4,4'-dimethoxytrityl cation, as measured by the intensity of its visible spectrum, so reducing its electrophilicity[130]. Using

non co-ordinating solvents such as acetonitrile, dichloromethane or nitromethane, the 4,4'-dimethoxytrityl carbocation was shown to have a strong absorption in visible spectrum suggesting that little, if any, donation of electron density from the solvent to the carbocation was occurring, potentially making the carbocation a stronger electrophile. Dimethylformamide and dimethyl sulfoxide (DMSO) both co-ordinated strongly with the 4,4'-dimethoxytrityl carbocation and quenched its colour. Basic conditions are however required. Addition of a co-ordinating base such as Hunig's base or triethylamine were found to quench the colour of the 4,4'-dimethoxytrityl carbocation but the use of the hindered 2,6-di-*tert*-butyl-4-methylpyridine (DBMP) was found to be satisfactory. It was also suggested that the rate of 4,4'-dimethoxytritylation of alcohol groups is a function of the concentration of the 4,4'-dimethoxytrityl carbocation. 4,4'-Dimethoxytrityl chloride in pyridine partially dissociates in an equilibrium process to form the 4,4'-dimethoxytrityl cation and pyridine hydrochloride. The 4,4'-dimethoxytrityl tetrafluoroborate salt, due to its ionic nature, should dissociate in solution thus providing maximum potential for 5'-OH 4,4'-dimethoxytritylation (Figure 2.14).

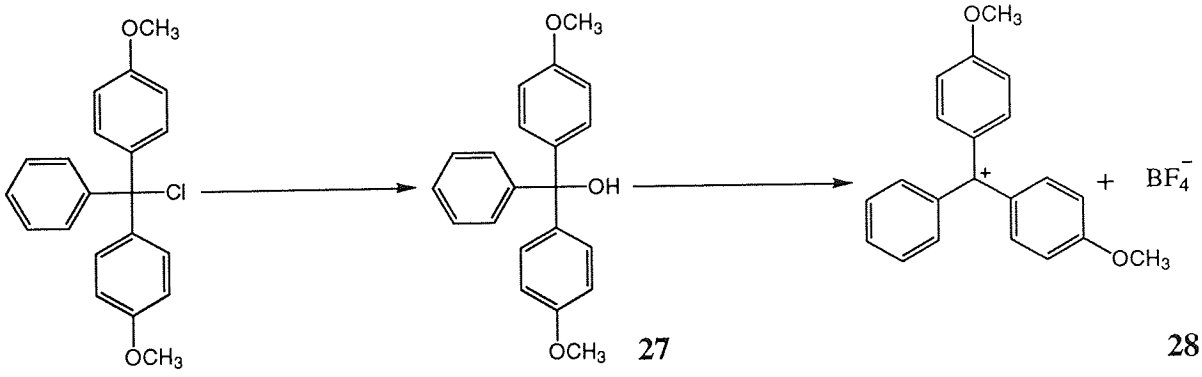


Figure 2.14 Synthesis of 4,4'-dimethoxytrityl tetrafluoroborate.

In accordance with the literature,[130] 4,4'-dimethoxytrityl tetrafluoroborate was made in 95% yield. However, 4,4'-dimethoxytritylation of **23** and **26** using 4,4'-dimethoxytrityl tetrafluoroborate (4 eq.) and DBMP (2 eq.) in various solvents remained unsatisfactory (Table 2.1). After refluxing the nucleoside with 4,4'-dimethoxytrityl tetrafluoroborate for 24

hours in nitromethane, the product **24** was isolated by flash chromatography in only 3% yield.

Reagent	Solvent	Conditions	TLC evidence?	Isolated?
Nucleoside 23:				
DMTrCl	Pyridine	Rt, 20 h	No	No
DMTrCl, DMAP	Pyridine	Rt, 72 h	Yes	No
DMTrCl, DMAP	DMF	75°C, 15 h	Yes	No
DMTrCl, Imidazole	Pyridine	Rt, 6 h	Yes	No
DMTrCl, Imidazole	Pyridine	75°C, 14 h	Yes	No
$DMTr^+BF_4^-$, DBMP	CH_3NO_2	Reflux, 24 h	Yes	Yes, 3%
$DMTr^+BF_4^-$, DBMP	CH_3CN	Reflux, 24 h	Yes	No
Nucleoside 26:				
DMTrCl	Pyridine	Rt, 20 h	No	No
DMTrCl, Imidazole	Pyridine	Rt, 22 h	No	No
$DMTr^+BF_4^-$, DBMP	CH_3NO_2	Rt, 42 h	Yes	No
$DMTr^+BF_4^-$, DBMP	CH_3CN	Reflux, 48 h	Yes	No
$DMTr^+BF_4^-$, DBMP	CH_2Cl_2	Reflux, 69 h	No	No

Table 2.1 Conditions and reagents used in the attempted synthesis of nucleosides **23** and **26**.

2.4.3 Alternative approaches to the synthesis of 24

An alternative approach was required to synthesise **24**. 2'-Deoxyuridine was 4,4-dimethoxytritylated in pyridine during 2.25 hours giving **4** in 88% yield following the standard protocol.[65] The 3'-hydroxyl was acetylated in pyridine giving **29** in 90% yield. However, all attempts to make the C4-triazolyl nucleoside using 4-Chlorophenyl phosphorodichloridate (2 eq.) in pyridine for C4 activation to make **30**, met with failure (Figure 2.15).

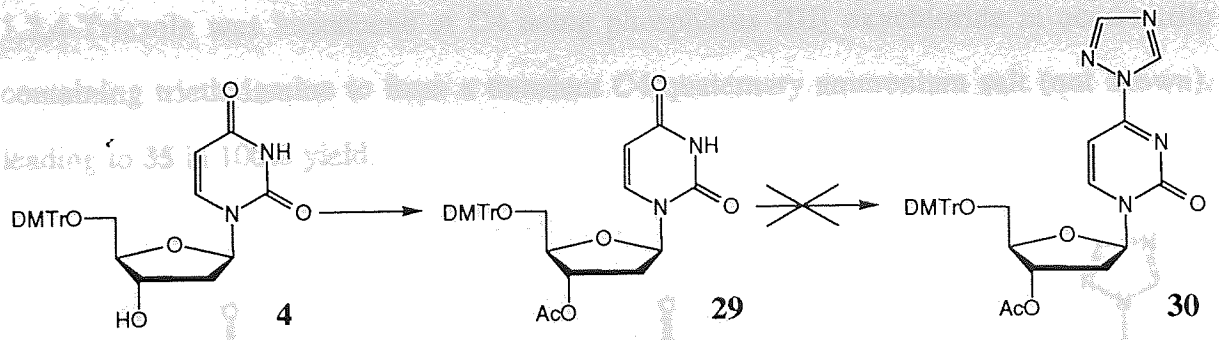


Figure 2.15 Attempted synthesis of the 5'-*O*-(4,4'-dimethoxytrityl)-C4-triazolyl pyrimidine **30**.

The failure of this reaction was surprising as the literature shows the successful synthesis of **32** from **31** where the 5'-hydroxyl of 2'-deoxyuridine was protected by 4,4'-dimethoxytrityl.[120] Nucleoside **31** was dissolved in pyridine to which 4-Chlorophenyl phosphorodichloridate (1.25 eq.) and 1,2,4-triazole (4.5 eq.) were added and stirred at 4°C for 4 days to give **32** in 78% yield. Nucleoside **32** was shown to undergo attack by amine nucleophiles at C4, giving products of the type **33** (Figure 2.16).

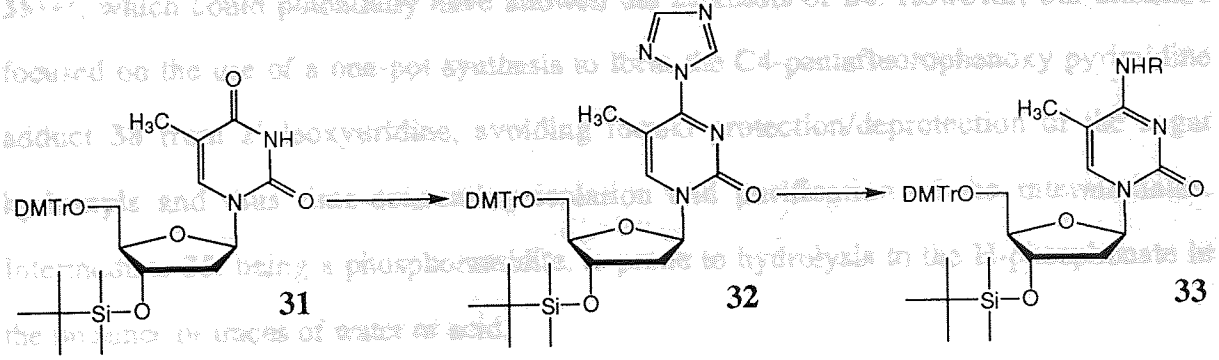


Figure 2.16 Literature synthesis of 5'-*O*-(4,4'-dimethoxytrityl)-C4 aralkyl-substituted pyrimidine nucleosides.

The literature also shows the synthesis of a C4 triazolyl nucleoside **35** which has been incorporated into oligonucleotides[127,131] using solid phase chemistry (Chapter 3). This allowed site-specific modification of oligonucleotides using amines. 2'-Deoxyuridine was 4,4'-dimethoxytritylated at the 5'-hydroxyl yielding **4**. The 3'-hydroxyl of **32** was phosphitylated using 2-cyanoethyl-*N*,*N*-di*iso*propyl chlorophosphoramidite yielding **34**.

1,2,4-Triazole was introduced at C4 using phosphorus (III) oxychloride in acetonitrile containing triethylamine to form a transient C4 quaternary ammonium salt (not shown), leading to **35** in 100% yield.

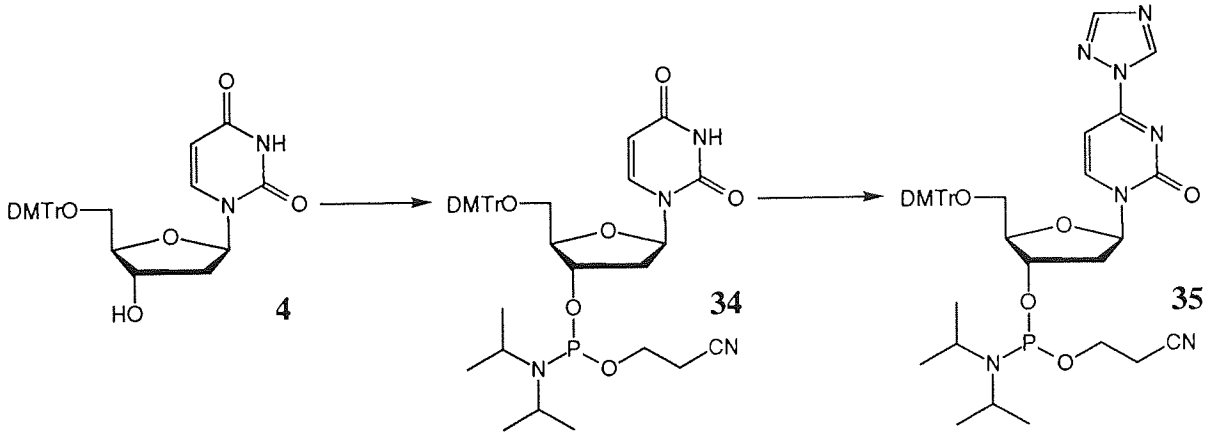


Figure 2.17 Literature synthesis of C4-triazolyl adduct **35** suitable for solid phase oligonucleotide synthesis.

The literature provided synthetic details for the preparation of two intermediates, **31**[131] and **35**[127], which could potentially have allowed the synthesis of **24**. However, our attention focused on the use of a one-pot synthesis to form the C4-pentafluorophenoxy pyrimidine adduct **38** from 2'-deoxyuridine, avoiding formal protection/deprotection of the sugar hydroxyls and thus time-consuming isolation and purification of the intermediates. Intermediate **35**, being a phosphoramidite, is prone to hydrolysis to the H-phosphonate in the presence of traces of water or acid.

Transient protection of the sugar hydroxyls of 2'-deoxyuridine is required to prevent side reactions whilst using electrophiles to activate the nucleoside at C4. Zhou *et al* (1986)[122] used trimethylsilyl chloride to transiently protect the 2'-hydroxyl of **12**. Deprotection of the 2'-hydroxyl was achieved using TBAF in the same reaction pot. Gao *et al* 1992, used trifluoroacetic anhydride to transiently protect the hydroxyl function of 2'-deoxyguanine. Two successful synthetic strategies were devised in this present work to introduce a

pentafluorophenoxy group at C4 of 2'-deoxyuridine, using either trimethylsilyl or trifluoroacetyl as transient protecting groups in one pot reactions.

2.4.4 One-pot synthesis of the C4-pentafluorophenoxy adduct 38 from 2'-deoxyuridine: method A

Application of the methodology of Gao *et al* 1992,[124] to introduce the pentafluorophenoxy moiety at C4 of 2'-deoxyuridine initially proved problematic. 2'-Deoxyuridine was dried by co-evaporation with pyridine and dissolved in pyridine. Trifluoroacetic anhydride (6 eq.) was added and the reaction mixture stirred for 5 hours to which pentafluorophenol was added (13 eq.). The reaction mixture was stirred for 24 h before being diluted with an equal volume of water. Extraction of the reaction mixture, using ethyl acetate, gave **38** in 9% yield with almost quantitative recovery of the remaining, unreacted 2'-deoxyuridine (Figure 2.18).

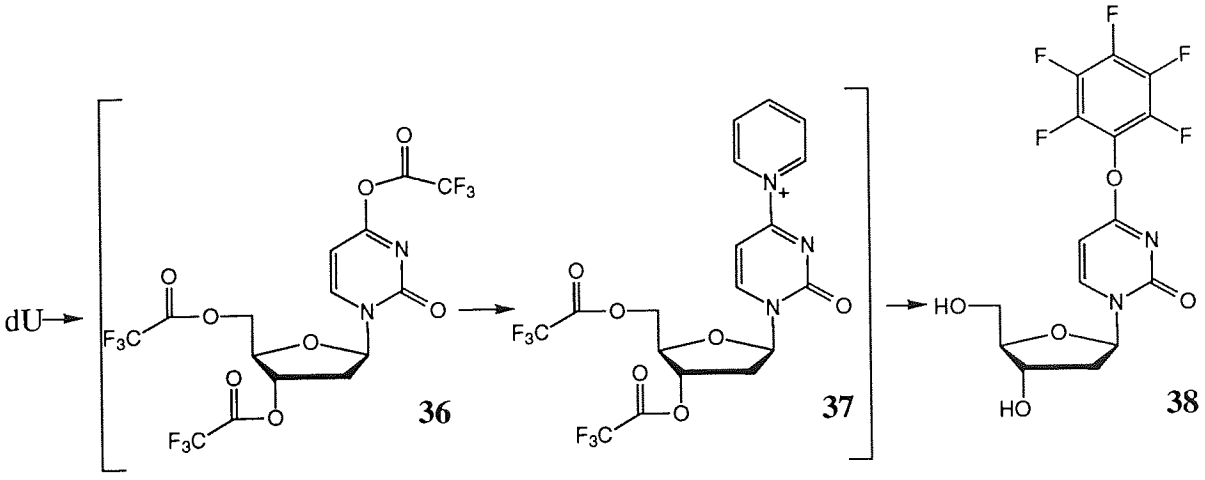


Figure 2.18 Synthesis of C4-pentafluorophenoxy adduct **38** using transient trifluoroacetyl sugar protection.

In an attempt to optimise the reaction conditions to synthesise **38** in an acceptable yield, the reaction was followed by TLC using ethyl acetate-methanol 9:1 as the mobile phase. After addition of trifluoroacetic anhydride to 2'-deoxyuridine in pyridine, a spot developed on the baseline during 5 hours although a spot which was attributed to 2'-deoxyuridine was also observed. Addition of pentafluorophenol showed a rapid re-formation of 2'-deoxyuridine. It

was not known if the spot on the baseline of the TLC plate was due to the C4 trifluoroacetyl ester **36** or the likely C4 pyridyl intermediate **37**. Attempts to resolve the spot on the base line by increasing the polarity of the mobile phase failed.

To investigate the formation of the pyridyl intermediate **37** from 2'-deoxyuridine, a ^1H NMR kinetic study was undertaken. 2'-Deoxyuridine was dissolved in d_5-pyridine in a NMR tube to which trifluoroacetic anhydride was added. The reaction was scanned 16 times after 10 minutes and then 16 times every 90 minutes. The resulting spectra were printed as a stack plot (Figure 2.19). After 1000 minutes the reaction had almost gone to completion. Consideration of the TLC and ^1H NMR evidence suggested that the formation of **36** was almost instantaneous while the formation of **37** was still incomplete after nearly 17 hours. The reaction was repeated on a preparative scale, allowing 24 hours before the addition of pentafluorophenol. TLC analysis showed formation of product **38** and indicated that all of the 2'-deoxyuridine had been consumed although other UV active spots were visible. Isolation of the product, following the literature procedure developed for the purine analogue **16**, gave **38** in 73% yield. Further attempts were made to improve the reaction yield of **38** by reducing the ratio of 2'-deoxyuridine to trifluoroacetic anhydride and pentafluorophenol. It was found that by using 4 equivalents of trifluoroacetic anhydride and 10 equivalents of pentafluorophenol, and isolating the product by flash chromatography, gave **38** as a white powder in acceptable 78% yield.

2.4.5 One-pot synthesis of the C4-pentafluorophenoxy adduct 38 from 2'-deoxyuridine: method B

A second method was developed in parallel for the synthesis of **38**. Transient protection of the sugar hydroxyl functions using trimethylsilyl ethers, enabled a one pot synthesis of **31** *via* **39**. Activation of C4, achieved using 4-Chlorophenyl phosphorodichloridate, was thought to facilitate the formation of the pyridyl intermediate **40** (Figure 2.20). 2'-Deoxyuridine was dried by co-evaporation with anhydrous pyridine, placed under argon and dissolved in pyridine to which trimethylsilyl chloride (2.2 eq.) was slowly added.

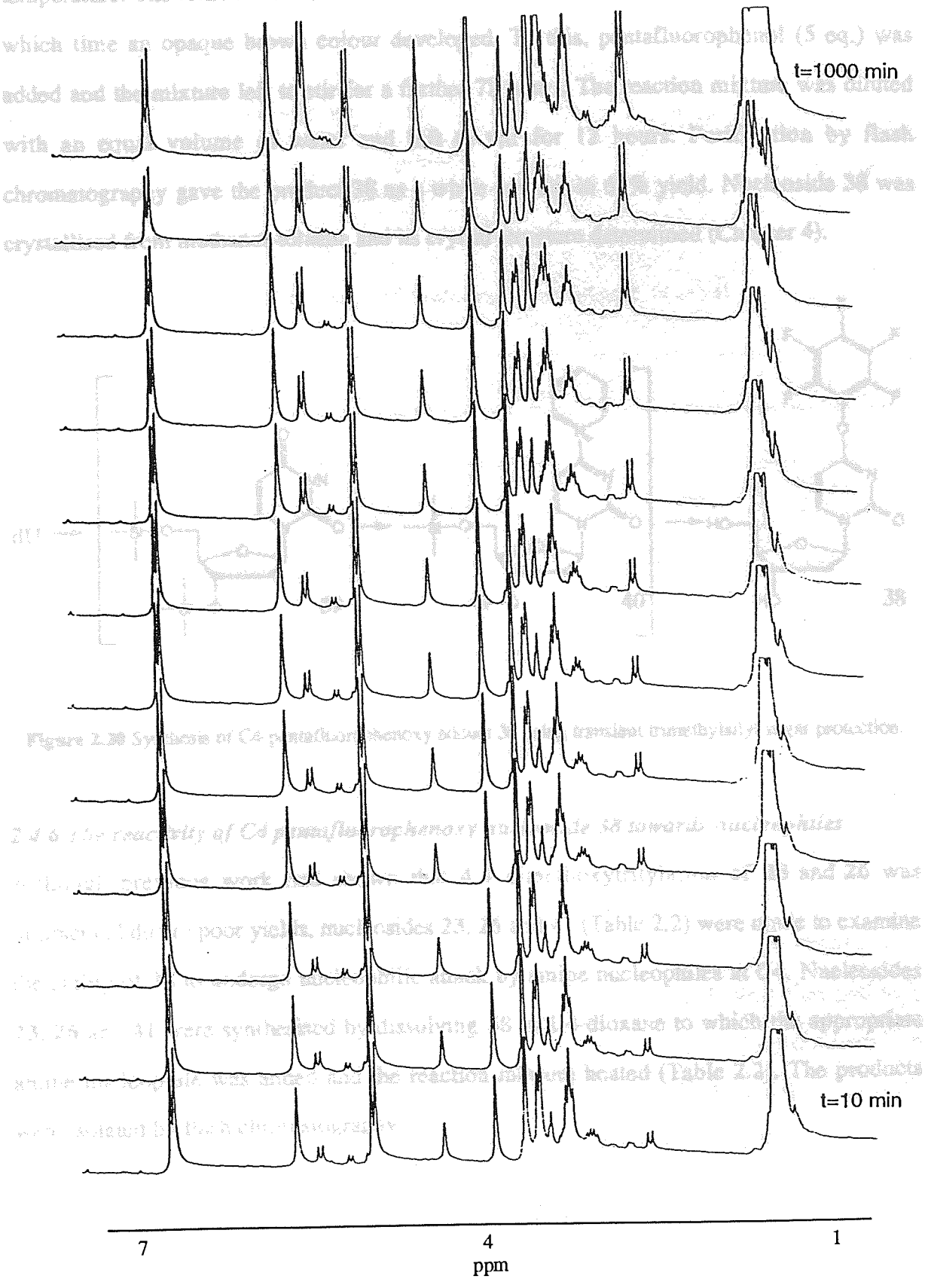


Figure 2.19 ^{1}H NMR kinetic study of the formation of the proposed intermediate **37**.

After 30 minutes the mixture was cooled to 0°C before 4-chlorophenyl phosphorodichloridate (2.5 eq.) was added and the reaction allowed to warm to room temperature. The reaction mixture was allowed to stir over a period over of 24 hours during which time an opaque brown colour developed. To this, pentafluorophenol (5 eq.) was added and the mixture left to stir for a further 72 hours. The reaction mixture was diluted with an equal volume of water and left to stir for 12 hours. Purification by flash chromatography gave the product **38** as a white powder in 65% yield. Nucleoside **38** was crystallised from methanol-toluene and its crystal structure determined (Chapter 4).

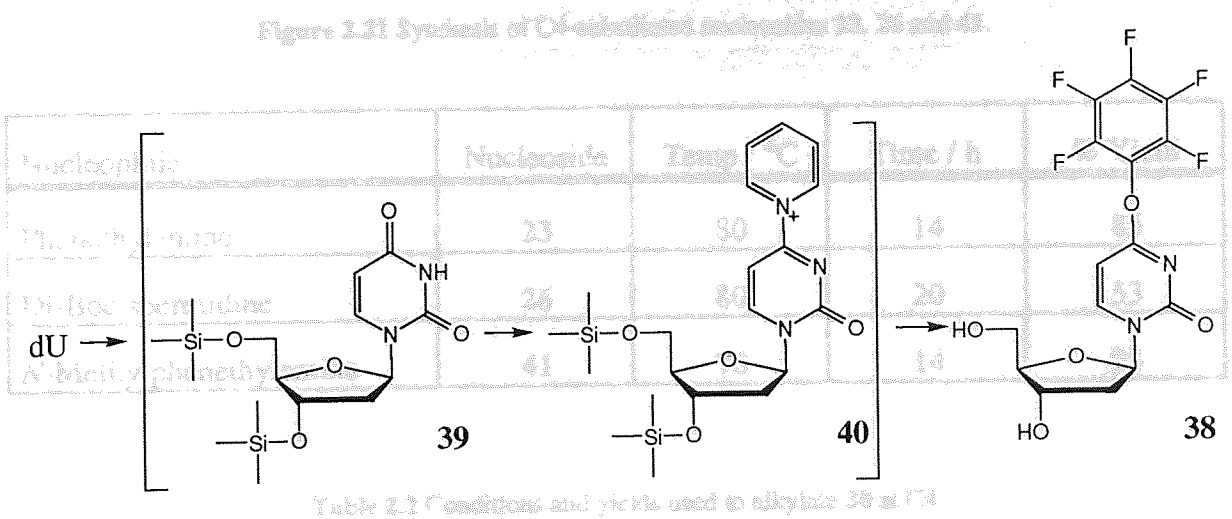


Figure 2.20 Synthesis of C4-pentafluorophenoxy adduct **38** using transient trimethylsilyl sugar protection.

2.4.6 The reactivity of C4 pentafluorophenoxy nucleoside 38 towards nucleophiles

Although previous work had shown that 4,4'-dimethoxytritylation of **23** and **26** was impractical due to poor yields, nucleosides **23**, **26** and **41** (Table 2.2) were made to examine the ability of **38** to undergo nucleophilic attack by amine nucleophiles at C4. Nucleosides **23**, **26** and **41** were synthesised by dissolving **38** in 1,4-dioxane to which the appropriate amine nucleophile was added and the reaction mixture heated (Table 2.2). The products were isolated by flash chromatography.

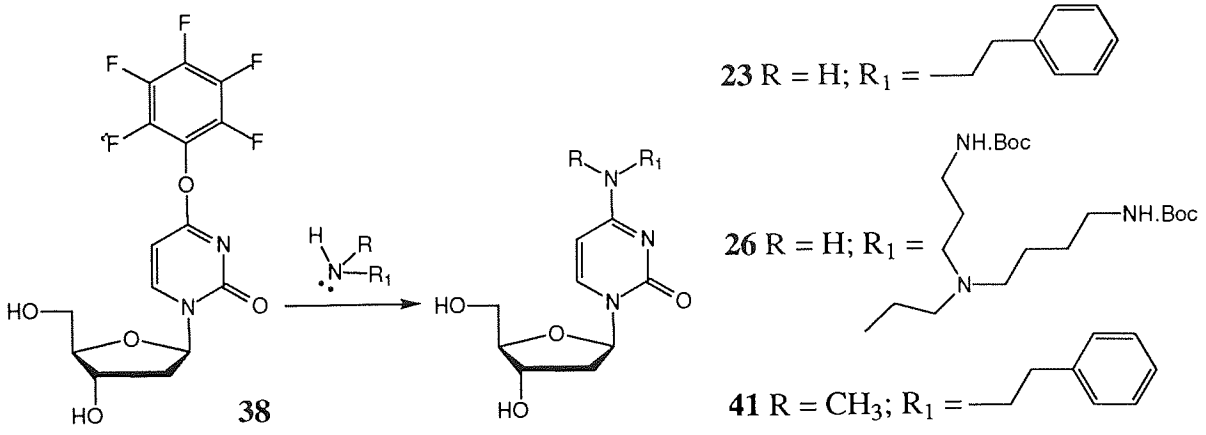


Figure 2.21 Synthesis of C4-substituted nucleosides **23**, **26** and **41**.

Nucleophile	Nucleoside	Temp / °C	Time / h	% Yield
Phenethylamine	**23**	80	14	85
Di-Boc spermidine	**26**	80	20	53
N-Methylphenethylamine	**41**	70	14	98

Table 2.2 Conditions and yields used to alkylate **38** at C4.

2.4.7 Observation of rotamers for N4 aralkylated cytosines

The ^{1}H and ^{13}C NMR spectra of **41** showed that two different structural isomers were present (Figure 2.22) unlike **23** where only a single isomer was apparent by NMR. The ^{1}H NMR spectrum of **41** at 20°C (Figure 2.22 (a)) showed sharp definition for the aromatic protons of the phenyl group but C5 and C6 at 7.90 ppm and 6.03 ppm were observed as two overlapping poorly resolved doublets. The signal for C1' at 6.17 ppm was observed as a broad singlet. In the ^{1}H NMR of **23** (not shown), C5 and C6 appear as sharp doublets and C1' is shown as a triplet. The ^{1}H NMR of **41** run at 30°C (Figure 2.22 (b)) shows C5 and C6 to be broad singlets at 7.88 ppm and 6.16 ppm while C1' appears as a well defined triplet at 6.16 ppm.

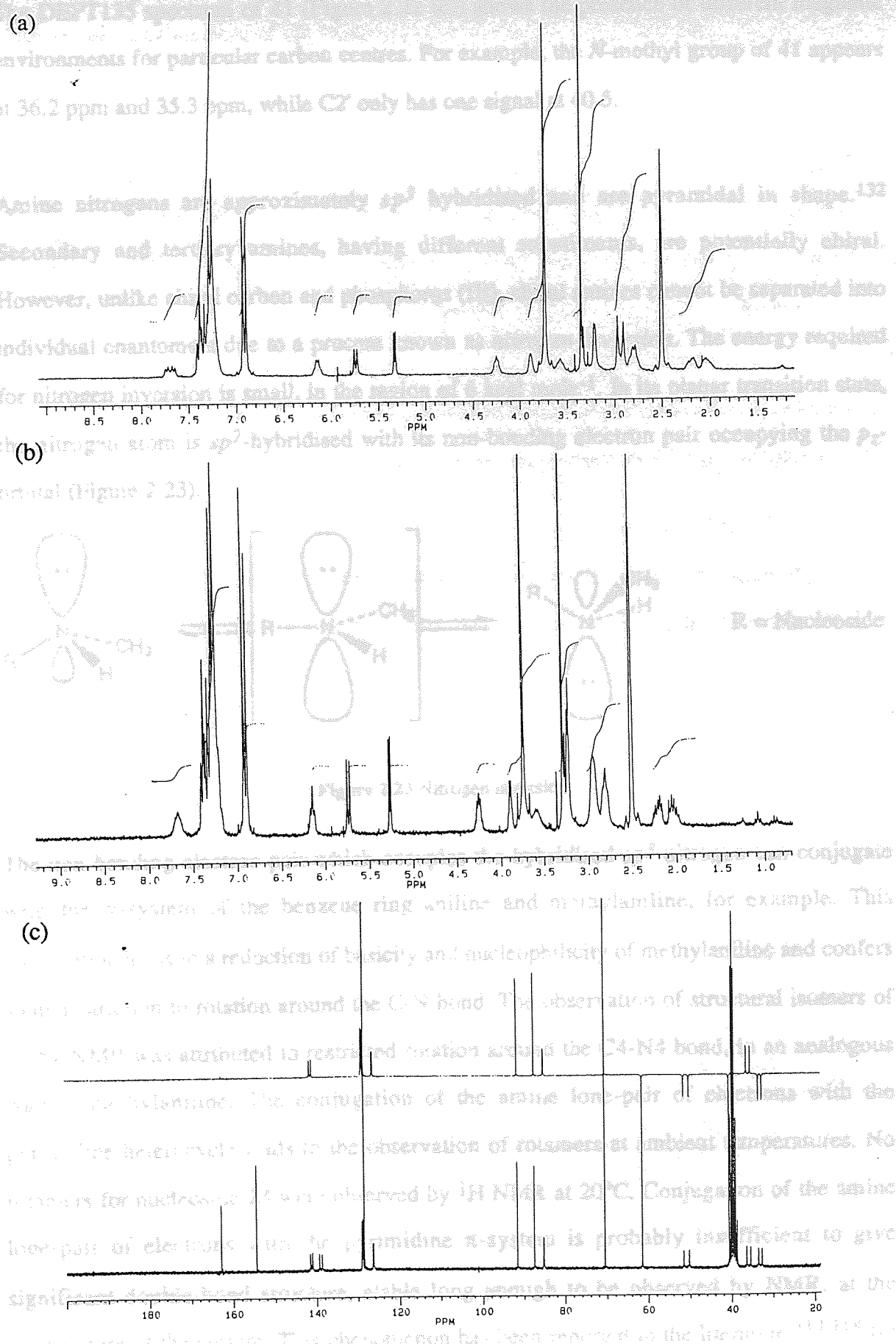


Figure 2.22 ^{1}H NMR of **41** at (a) 20°C; (b) 35°C and (c) a carbon/DEPT 135 overlay of **41** at 20°C.

The DEPT135 spectrum of **41** (Figure 2.22 (c)) shows the presence of different magnetic environments for particular carbon centres. For example, the *N*-methyl group of **41** appears at 36.2 ppm and 35.3 ppm, while C2' only has one signal at 40.5.

Amine nitrogens are approximately sp^3 hybridised and are pyramidal in shape.[132] Secondary and tertiary amines, having different substituents, are potentially chiral. However, unlike chiral carbon and phosphorus (III), chiral amines cannot be separated into individual enantomers due to a process known as *nitrogen inversion*. The energy required for nitrogen inversion is small, in the region of 6 kcal mole^{-1}. In its planar transition state, the nitrogen atom is sp^2-hybridised with its non-bonding electron pair occupying the p_z-orbital (Figure 2.23).

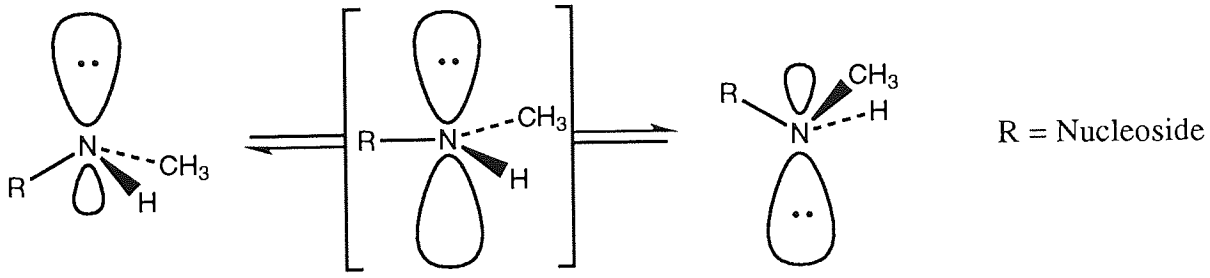


Figure 2.23 Nitrogen inversion.

The non-bonding electron pair which occupies the hybridised sp^3 nitrogen can conjugate with the π-system of the benzene ring aniline and methylaniline, for example. This conjugation leads to a reduction of basicity and nucleophilicity of methylaniline and confers some restriction to rotation around the C-N bond. The observation of structural isomers of **41** by NMR was attributed to restricted rotation around the C4-N4 bond, in an analogous way to methylaniline. The conjugation of the amine lone-pair of electrons with the pyrimidine heterocycle leads to the observation of rotamers at ambient temperatures. No rotamers for nucleoside **23** were observed by ^{1}H NMR at 20°C. Conjugation of the amine lone-pair of electrons with the pyrimidine π-system is probably insufficient to give significant double-bond structure, stable long enough to be observed by NMR, at the temperature of the system. This phenomenon has been reported in the literature,[117,118] for

example, N6-methylation of adenosine causes restriction around the N6-methyl bond and alters the electronic nature of the base.

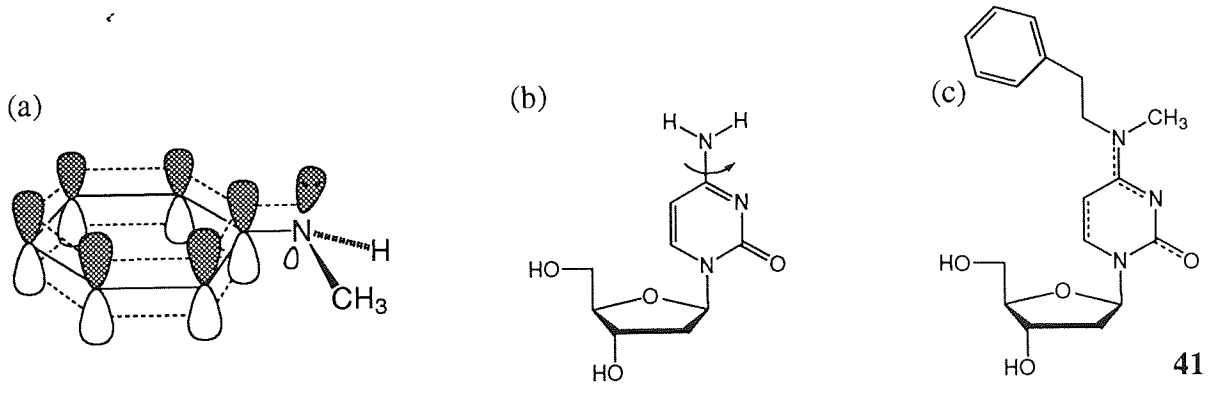


Figure 2.24 Schematic representation of (a) conjugation of the amine lone pair of electrons and an aromatic π system; (b) free rotation of the cytosine amine function at room temperature and (c) restricted rotation of **41**.

2.4.8 4,4'-Dimethoxytritylation of C4 pentafluorophenoxy pyrimidine adduct 38

Nucleoside **38** was 4',4-dimethoxytritylated using 4',4-dimethoxytrityl chloride in pyridine at room temperature during 2 hours. Purification by flash chromatography gave **42** as a foam in 85% yield.

2.4.9 Reactivity of C4 pentafluorophenoxy pyrimidine 42 towards nucleophiles

Displacement of the C4 triazolyl group of **42** by amine nucleophiles phenethylamine, *N*-methylphenethylamine and spermidine yielded **24**, **43** and **44** in good yield (Table 2.3). Rotamers of **44** were observed by NMR analysis.

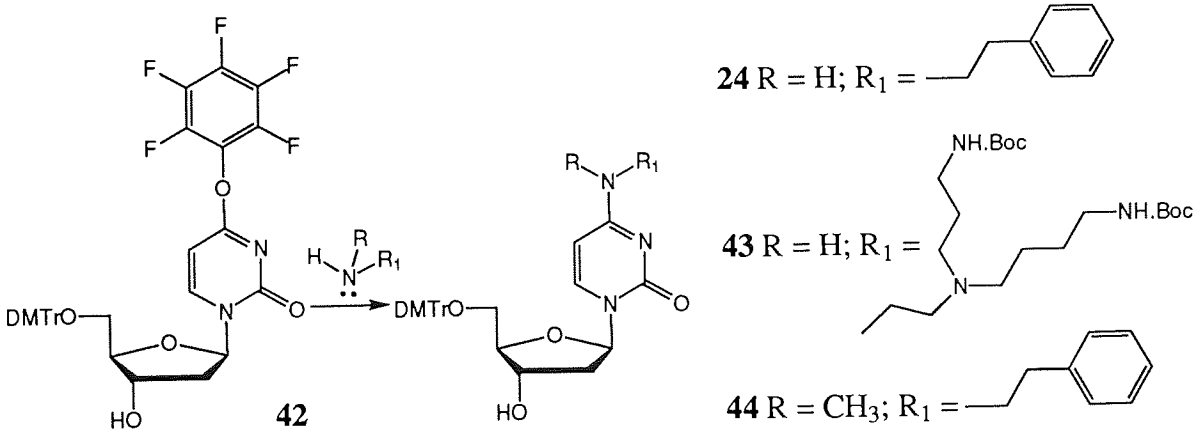


Figure 2.25 Synthesis of 5'-*O*-(4,4'-dimethoxytrityl)-C4-substituted nucleosides **24**, **43** and **44**.

Nucleophile:	Nucleoside	Temp / °C	Time / h	% Yield
Phenethylamine	**24**	70	17	95
Spermidine	**43**	90	36	71
N-Methylphenethylamine	**44**	70	14	81

Table 2.3 Conditions and yields used to aralkylate **42** at C4.

2.4.10 Phosphitylation of C4 aralkyl-substituted 2'-deoxycytosines

To fulfil the chemical requirements for solid phase oligonucleotide synthesis, the 3'-[(2-cyanoethyl)-*N,N*-di*iso*propyl]phosphoramidite derivatives of the 4,4'-dimethoxytrityl nucleosides are required[65] (Chapter 3). Phosphoramidites are susceptible to H-phosphonate formation in the presence of water. Prior to the phosphitylation reaction, the 4,4'-dimethoxytrityl nucleosides **24**, **43** and **44** were dried at 40°C under vacuum in the presence of P_2O_5. All reactions were carried out using anhydrous THF as the solvent with an anhydrous base (2eq.) and 3'-[(2-cyanoethyl)-*N,N*-di*iso*propyl]phosphoramidite chloride as the phosphitylating agent.

Nucleoside	Product	Base	Eq. of phosphoramidite	Form	Yield %
42	**45**	triethylamine	2	Foam	100
24	**46**	triethylamine	1.1	Foam	87
43	**47**	di*iso*propylethylamine	2	Oil	87
44	**48**	di*iso*propylethylamine	1.5	Oil	30*

Table 2.4 Conditions and reagents used for 3'-*O*-phosphitylation of 5'-*O*-4,4'-dimethoxytrityl-nucleosides; * as judged by ^{31}P NMR analysis.

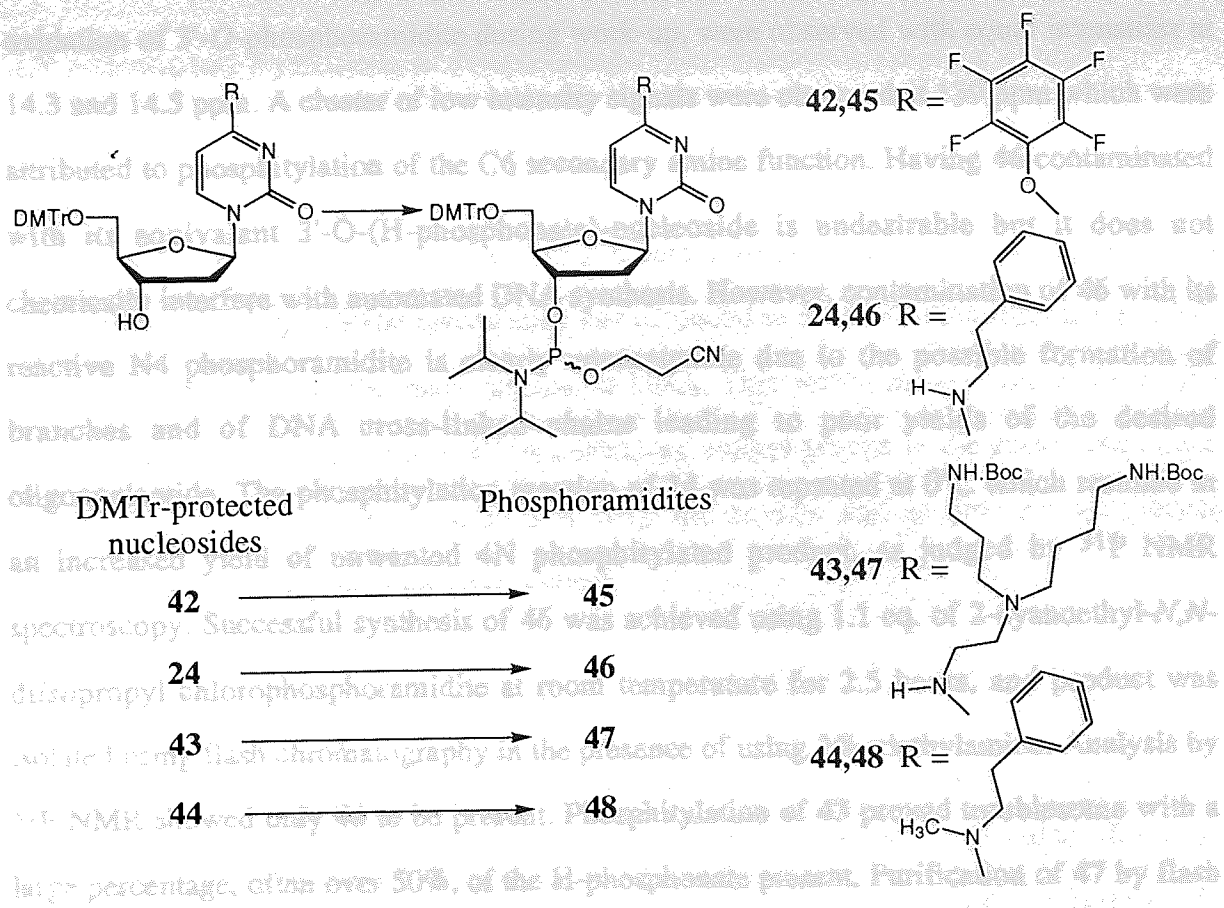


Figure 2.26 Phosphitylation of 5'-*O*-4,4'-dimethoxytrityl-nucleosides for solid phase oligonucleotide synthesis.

There are two potential phosphitylation sites on **24** and **43**; the 3'-hydroxyl and the N4 amino group. The exocyclic amino functions of purine and pyrimidine nucleosides require masking to prevent side reactions with the incoming phosphoramidite. Acylation of the N4-amine of 2'-deoxyadenosine and 2'-deoxycytosine with benzoyl chloride forms an unreactive amide function at C4. Initial attempts to phosphitylate **24** used 3'-[(2-cyanoethyl)-*N,N*-di*iso*propyl]phosphoramidite chloride (2 eq.). The crude product was filtered through cotton wool, concentrated under vacuum, suspended in d_6-DMSO and analysed by ^{31}P NMR. Three sets of peaks were observed. The major set, 50% by peak heights, were observed at 147.6 and 148.1 ppm and were assigned as the distereoisomers of the 3'-*O*-phosphoramidite of which the signal at 147.6 ppm was in excess by 2:1. H-Phosphonate, formed either by phosphorylation with oxidised phosphitylating agent or

oxidation of 3'-*O*-phosphoramidite during work-up, were observed with equal intensities at 14.3 and 14.5 ppm. A cluster of low intensity signals were observed at 139 ppm which were attributed to phosphitylation of the C6 secondary amine function. Having **46** contaminated with its equivalent 3'-O-(H-phosphonate)-nucleoside is undesirable but it does not chemically interfere with automated DNA synthesis. However, contamination of **46** with its reactive N4 phosphoramidite is clearly unacceptable due to the possible formation of branches and of DNA cross-linked chains leading to poor yields of the desired oligonucleotide. The phosphitylation reaction of **24** was repeated at 0°C which resulted in an increased yield of unwanted 4N phosphitylated product, as judged by ^{31}P NMR spectroscopy. Successful synthesis of **46** was achieved using 1.1 eq. of 2-cyanoethyl-*N,N*-di*iso*propyl chlorophosphoramidite at room temperature for 2.5 hours, and product was isolated using flash chromatography in the presence of using 2% triethylamine. Analysis by ^{31}P NMR showed only **46** to be present. Phosphitylation of **43** proved troublesome with a large percentage, often over 50%, of the H-phosphonate present. Purification of **47** by flash chromatography required relatively polar conditions (ethyl acetate-methanol 4:1) buffered with 1% triethylamine, and could not be isolated without major contamination by its equivalent H-phosphonate. Nucleoside **43** could not be characterised satisfactorily. Nucleosides **42** and **44** were phosphitylated using 2 equivalents of 2-cyanoethyl-*N,N*-di*iso*propyl chlorophosphoramidite as only one reactive site was available for phosphitylation.

Nucleosides **46** and **48** were incorporated into oligonucleotides to measure the effect of the N4 modified cytosines in terms of thermal stability (Chapter 3). Nucleoside **45** was incorporated into an oligonucleotide to test its susceptibility to nucleophilic attack by amines post-oligonucleotide assembly.

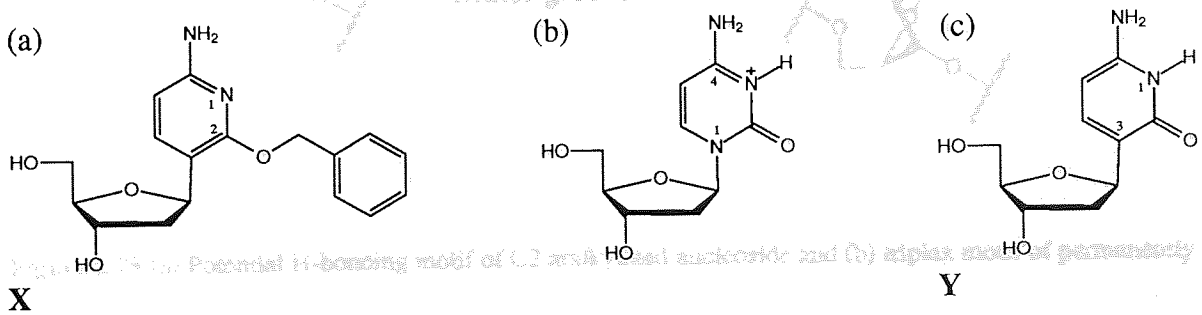
2.5 Attempted Synthesis of Nucleosides with Minor Groove Ligands

2.5.1 Rationale

Our interest in introducing aralkyl substituents into the major groove of DNA by aralkylating C4 of pyrimidine nucleosides was extended to attempt to investigate the effect of aralkyl substituent in the minor groove of DNA. This would allow comparison of the thermodynamic characteristics of DNA containing aralkyl groups in the major and minor groove. In B-DNA, the minor groove is deep and narrow and so presents a different chemical, steric and electrostatic environment.

2.5.2 Literature review, design and applications

The chosen method for introducing aralkyl groups into the minor groove of duplex DNA required aralkylation at C2 of a pyrimidine nucleoside. A C2 aralkyl nucleoside which has the required 'cytosine' H-bonding motif requires a tetravalent atom at C3 of the heterocycle (Figure 2.27 (a) and (b)). A pyridine ring, for example, has the correct structural requirement for the heterocyclic component of the nucleoside. Figure 2.27 (a) shows the synthetic target **X** for this work and the numbering system of the pyridine heterocycle.



X **Y**

Figure 2.27 (a) Synthetic target; (b) protonated 2'-deoxycytosine cation and (c) permanently 'protonated' 2'-deoxycytosine analogue.

Synthesis of **X** could have other potential applications. Hydrogenolysis of **X** could potentially yield **Y**, a permanently 'protonated' 2'-deoxycytosine analogue, which could

have applications in the antigene arena. Cytosine can only form a DNA triplex at acidic pH as N3 of the pyrimidine heterocycle required protonation to achieve the required H-bond motif[133] (Figure 2.7 (b)). The 'permanently protonated' cytosine analogue **Y** could potentially be used to form triplexes at non-acidic pH (Figure 2.27 (c)). Figure 2.28 (a) illustrates how the synthetic target **X** could H-bond with guanine whilst delivering the C2 aralkyl substituent into the minor groove of a DNA duplex. Figure 2.28 (b) illustrates how the 'permanently protonated' cytosine analogue **Y** could form a triplex complex with its target DNA duplex.

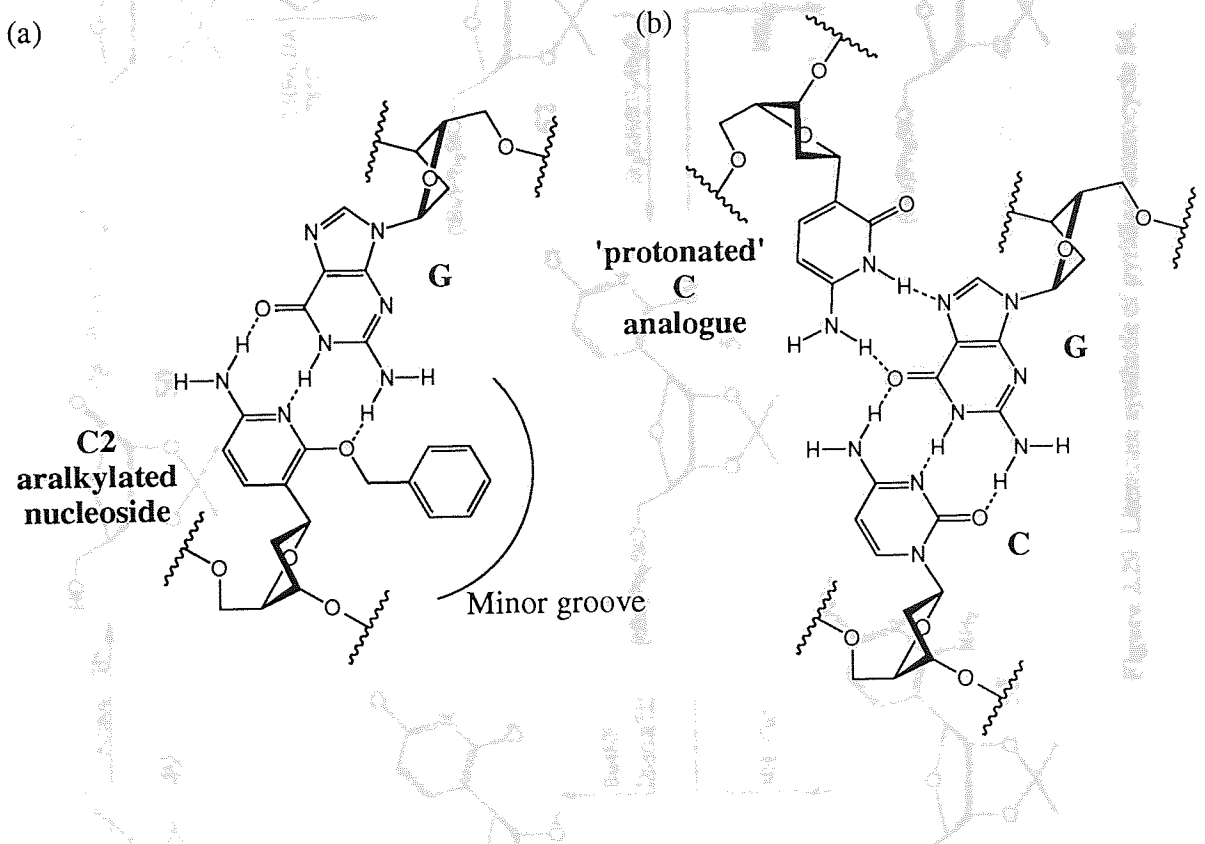


Figure 2.28 (a) Potential H-bonding motif of C2 aralkylated nucleoside and (b) triplex motif of permanently protonated cytosine.

Piccirilli *et al* (1991)[134] synthesised ribonucleosides **54**, **55** and **56** which could potentially be aralkylated at C2 (Figure 2.29). Nucleoside **52** was synthesised in the ratio α-D/β-D 1:7. Entry into the 2'-deoxynucleoside series was achieved by deoxygenation at C2' (Figure 2.30).

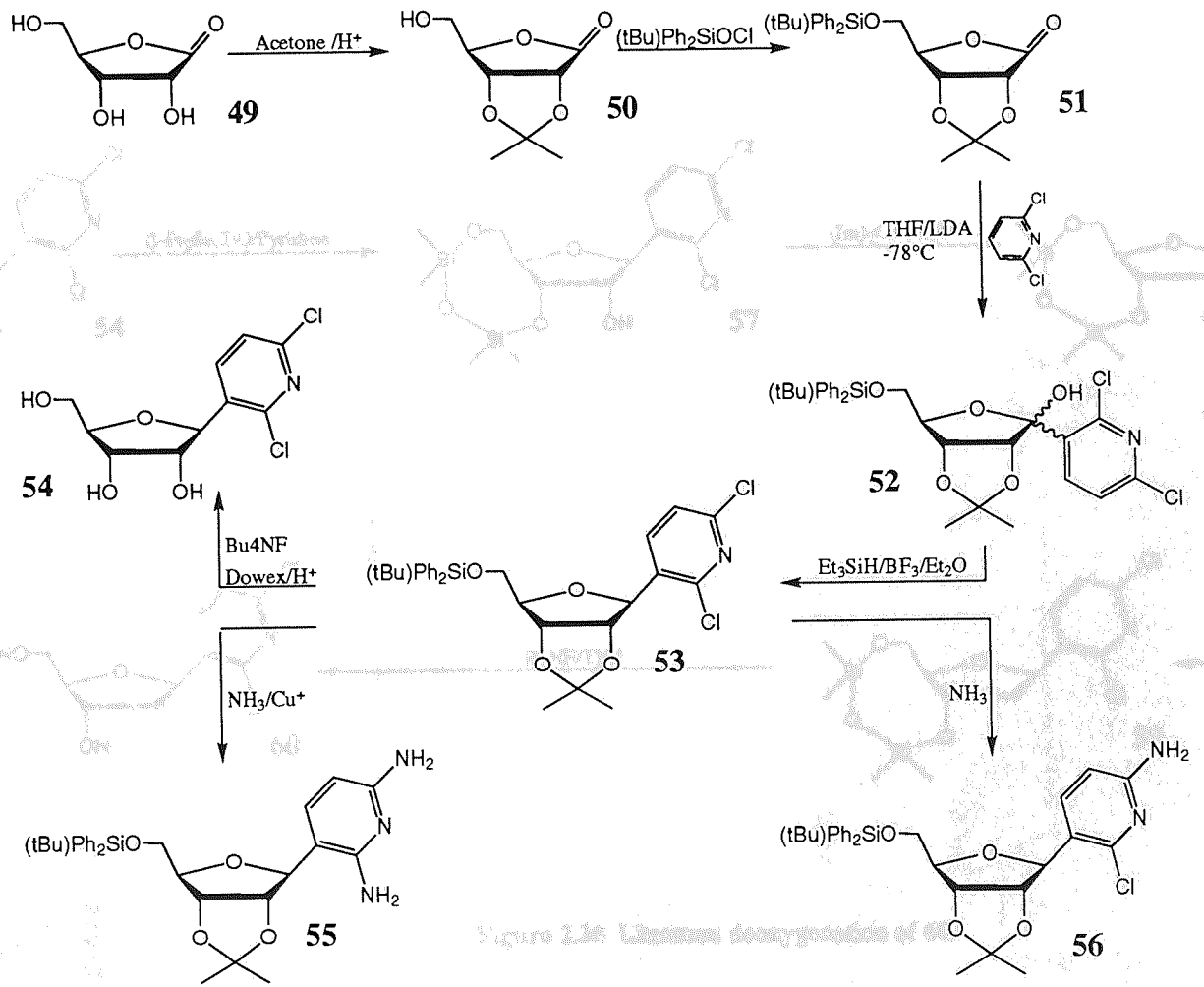


Figure 2.29 Literature synthesis of pyridine heterocycle **54**.

Figure 2.30 Literature deoxygenation of **60**.

A more direct route to C2 aralkyl nucleosides was sought. 2-Deoxy-D-α-chloro-ribose has been successfully used to make the dichloropurine nucleoside **62**[135] (Figure 2.31). It was considered that the use of this chloro sugar **61** to make the C-C bonded nucleoside **67** would reduce the number of structural isomeric by-products; the β-isomers **67** and **68** with respect to the sugar being the major anticipated products.

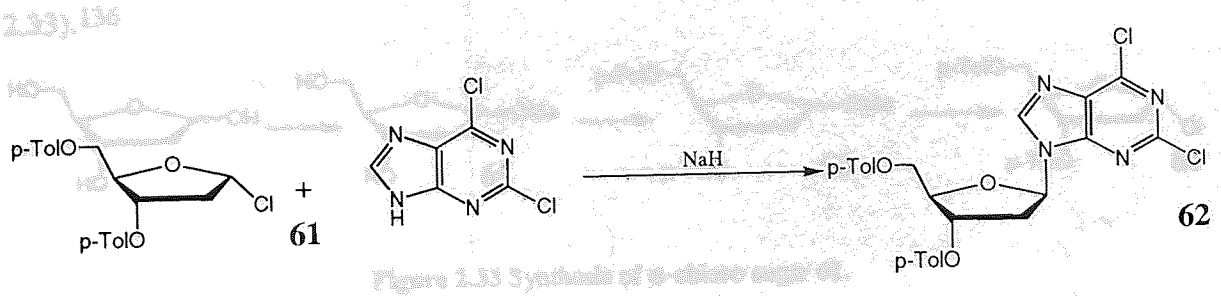


Figure 2.31 Synthesis of the 2,6-dichloropurine nucleoside **62**.

2.5.3 Synthetic work

The starting material for the synthesis of the pyridine heterocycle was 6-chloro-2-pyridol which was dispersed in toluene containing silver carbonate (1.5 eq.) and heated for one hour at 110°C before being cooled to 70°C to which benzyl bromide (2 eq.) was then added. The reaction mixture was stirred overnight at 110°C before the solid was filtered off through celite and the filtrate was concentrated under vacuum and purified by flash chromatography yielding the product **63** as a colourless viscous liquid in 63% yield (Figure 2.32). Before attempting to couple heterocycle **63** to α-chloro sugar **61**, the conditions necessary to aminate **63** were investigated (Figure 2.32).

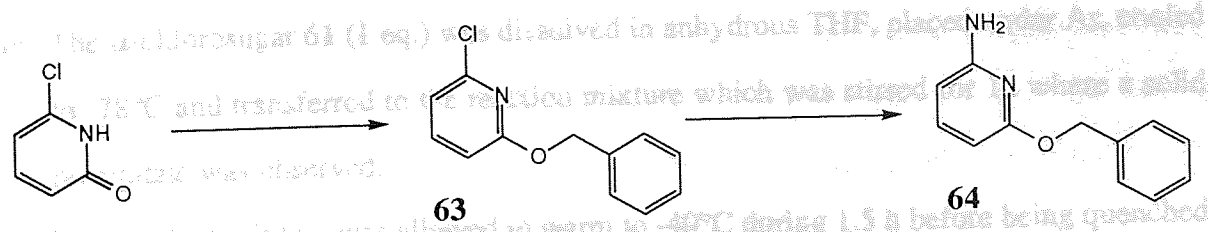


Figure 2.32 Synthesis of pyridine heterocycle **64**.

Heterocycle **63** was successfully aminated at C6 by dissolving **63** in saturated methanolic ammonia with Cu(I)Cl (2eq.) and heating in a bomb at 150°C for 15h. Heterocycle **64** was purified by flash chromatography and was characterised by ^{1}H-NMR.

The α-chlorosugar was synthesised in good yield according to the literature (Figure 2.33).[136]

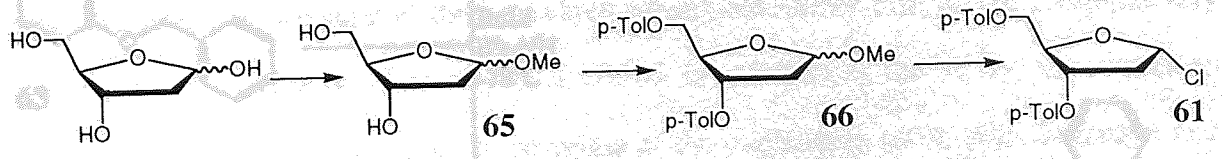


Figure 2.33 Synthesis of α-chloro sugar **61**.

Attempts to couple α-chlorosugar **61** with heterocycle **63** using LDA to form a carbanion of **63** met with failure (Figure 2.34).

The general procedure employed for the attempted synthesis of nucleosides **67** and **68** was as follows:

i. Di*iso*propylamine (1.1 eq.) was dissolved in anhydrous tetrahydrofuran (THF), placed under Ar and cooled to 0°C.

ii. Butyl lithium (1 eq.) was added to the solution which was stirred at 0°C for 0.5h before being cooled to -78°C.

iii. The pyridine derivative **63** (1 eq.) was dissolved in anhydrous THF, placed under Ar, cooled to -78°C and transferred to the reaction mixture which was stirred for 1h.

iv. The α-chlorosugar **61** (1 eq.) was dissolved in anhydrous THF, placed under Ar, cooled to -78°C and transferred to the reaction mixture which was stirred for 1h where a solid precipitate was observed.

v. The reaction mixture was allowed to warm to -40°C during 1.5 h before being quenched with water.

vi. The reaction mixture was concentrated under vacuum and analysed by TLC.

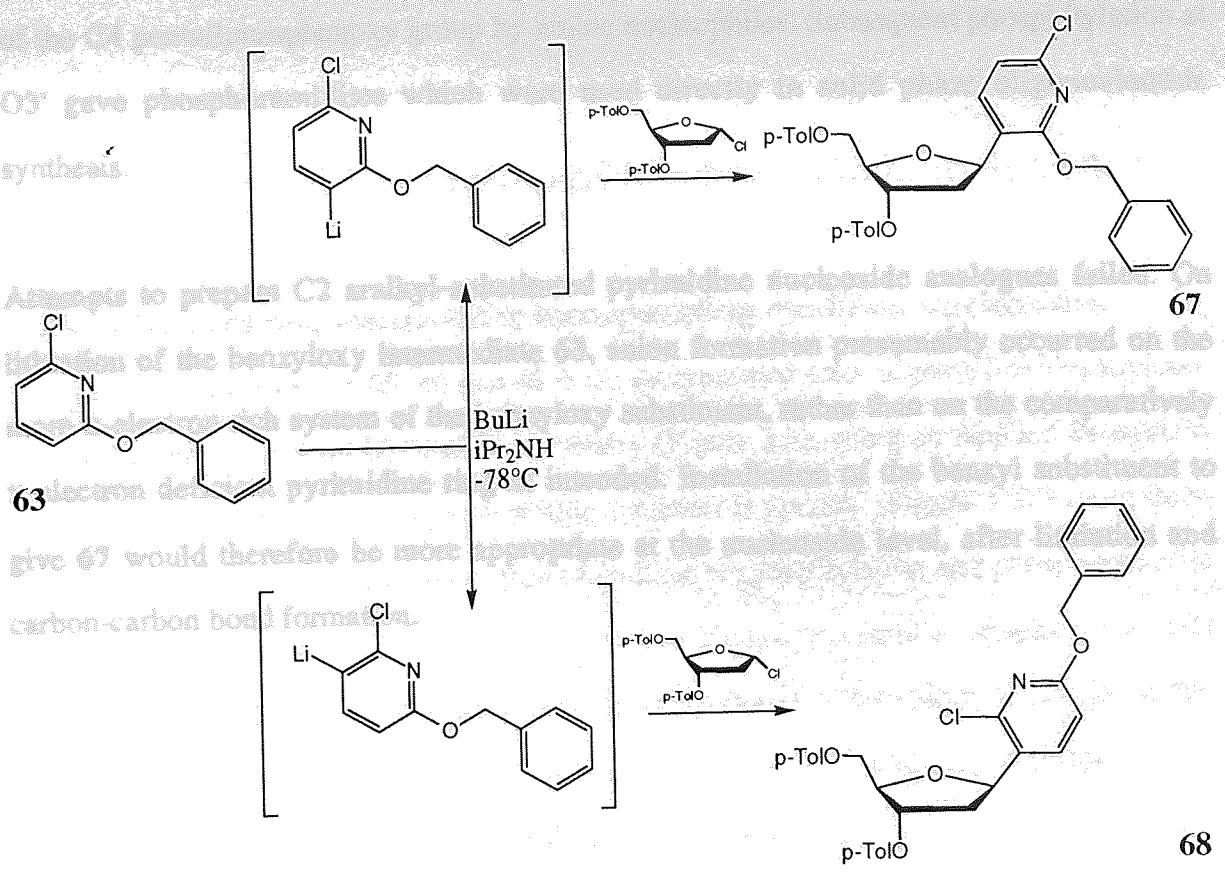


Figure 2.34 Attempted synthesis of nucleosides **67** and **68**.

TLC analysis showed consumption of α-chlorosugar **61** and heterocycle **63**. A number for UV active spots were seen and isolated but ^{1}H-NMR analysis showed them not to be nucleosides **67** and **68**. Compounds isolated from the reaction were not identified.

2.6 Conclusion

Efficient and direct methods have been successfully developed for the synthesis of N4 aralkyl-substituted pyrimidine nucleosides for incorporation into oligonucleotides. Pyrimidine nucleoside **38**, with its displaceable pentafluorophenoxy group at C4, can be synthesised directly in one-pot from 2'-deoxyuridine using *in situ* protection of the 3'- and 5'-hydroxyls. 4,4'-Dimethoxytritylation of C4 phenethyl-substituted nucleoside **23** was unsatisfactory but the use of **38**, which underwent 4,4'-dimethoxytritylation in good yield, allowed synthesis of 4,4'-dimethoxytrityl-protected N4-aralkyl nucleosides via displacement

of the C4 pentafluorophenoxy group by amine nucleophiles. Subsequent phosphitylation at O3' gave phosphoramidites which were used directly in solid phase oligonucleotide synthesis.

Attempts to prepare C2 aralkyl-substituted pyrimidine nucleoside analogues failed. On lithiation of the benzyloxy intermediate **63**, anion formation presumably occurred on the more π-electron rich system of the benzyloxy substituent, rather than on the comparatively π-electron deficient pyrimidine ring as intended. Installation of the benzyl substituent to give **67** would therefore be more appropriate at the nucleoside level, after lithiation and carbon-carbon bond formation.

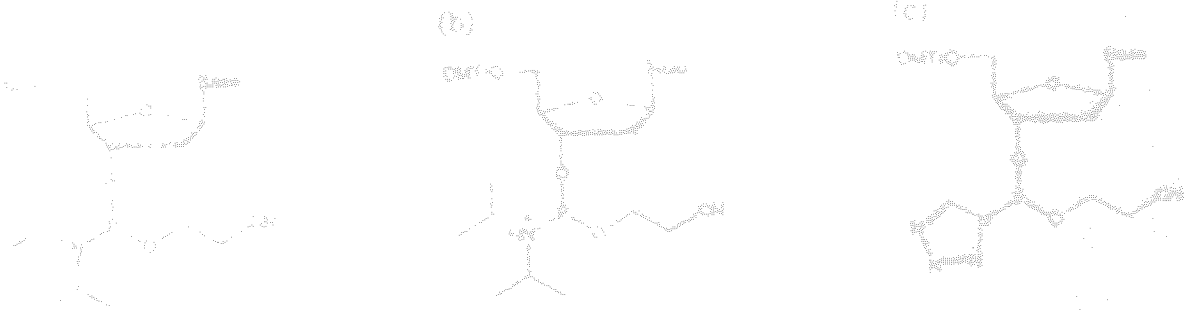


Figure 3.1 (a) The phosphoramidite nucleoside, (b) the protonated *N,N*-diisopropyl amine group and (c) the tetrazolyl leaving group.

Chapter 3

Synthesis and Evaluation of Modified Oligonucleotides

3.1 Synthesis of oligonucleotides incorporating modified nucleosides

The protected nucleosides **45**, **46** and **48** were incorporated into oligonucleotides utilising standard solid phase phosphoramidite chemistry (Figure 3.2), using an Applied Biosystem 392 synthesiser. Solid phase oligonucleotide synthesis is cyclical (Figure 3.2) having three main phases:[65] (**I**) nucleoside preparation (4,4'-dimethoxydetritylation and phosphoramidite activation); (**II**) P-O bond formation *via* 5'-OH to 3'-phosphoramidite coupling and (**III**) P(III) to P(V) oxidation. Once oligonucleotide synthesis is complete, cleavage of the oligonucleotide from the Controlled Glass Support (CPG) and deprotection of the oligonucleotide is achieved with saturated aqueous ammonia.

The synthetic cycle starts by activating the incoming nucleoside with tetrazole. The exact mechanism by which the phosphoramidite is activated is not clear but is thought to proceed in one of two ways (Figure 3.1), either (**i**) by protonating the *N,N*-di*iso*propyl amine group thus making it a better leaving group or (**ii**) by the forming an activated tetrazolyl-phosphoramidite intermediate by elimination of the *N,N*-di*iso*propyl amine group.[137]

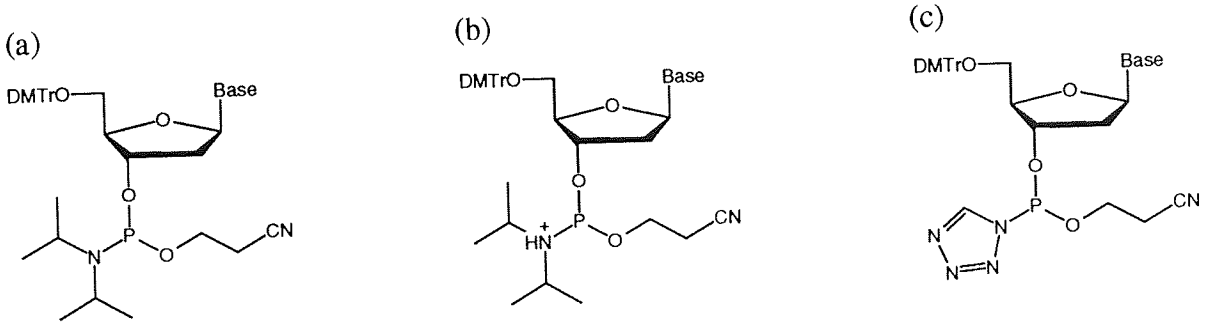


Figure 3.1 (a) The phosphoramidite nucleoside; (b) the protonated *N,N*-di*iso*propyl amine group and (c) the tetrazolyl leaving group.

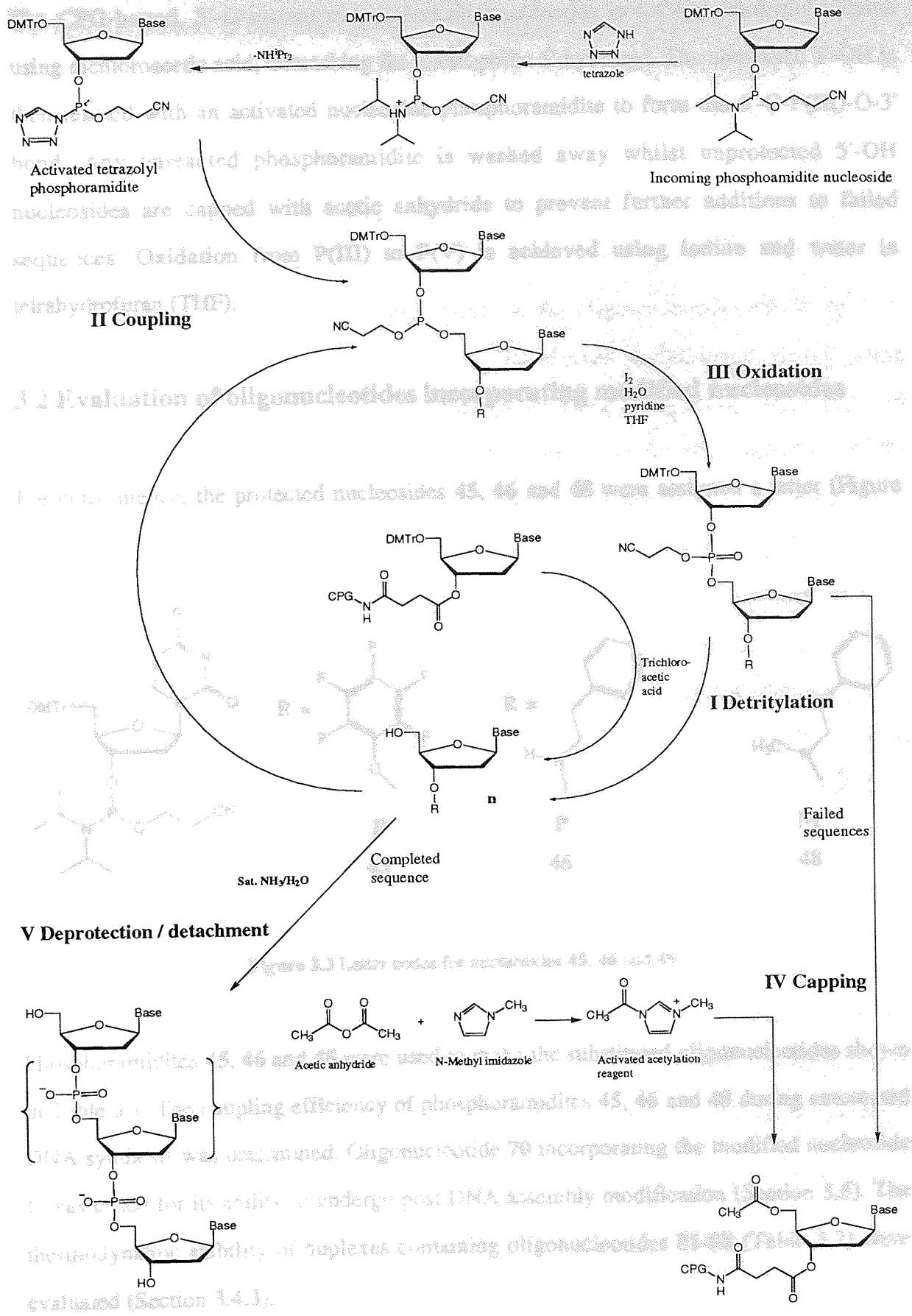
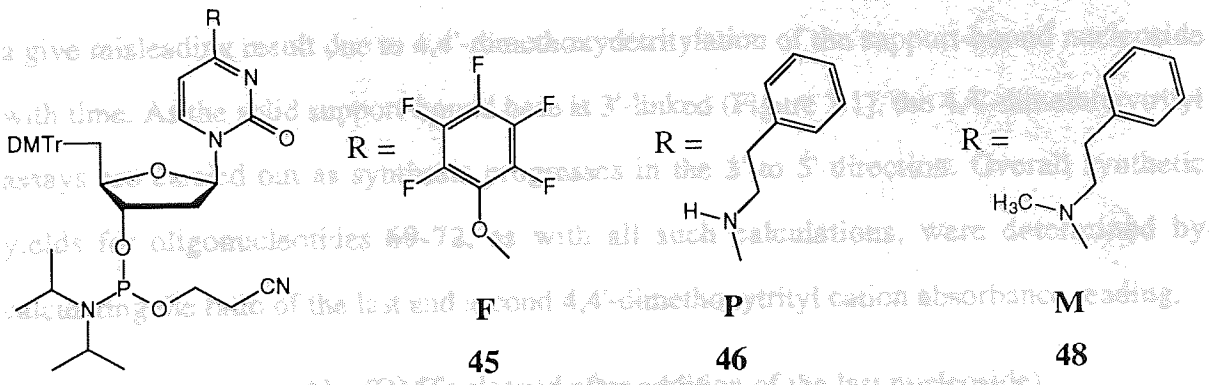


Figure 3.2 Solid phase oligonucleotide synthesis.

The CPG-bound, 3'-*O*-succinamide-linked oligonucleotide is 4,4'-dimethoxydetritylated using dichloroacetic acid, unmasking the nucleophilic 5'-hydroxyl. The unmasked 5'-OH is then reacted with an activated nucleoside phosphoramidite to form the 5'-O-P(III)-O-3' bond. Any unreacted phosphoramidite is washed away whilst unprotected 5'-OH nucleosides are capped with acetic anhydride to prevent further additions to failed sequences. Oxidation from P(III) to P(V) is achieved using iodine and water in tetrahydrofuran (THF).

3.2 Evaluation of oligonucleotides incorporating modified nucleosides

For convenience, the protected nucleosides **45**, **46** and **48** were assigned a letter (Figure 3.3).



F **P** **M**

45 **46** **48**

Figure 3.3 Letter codes for nucleosides **45**, **46** and **48**.

Phosphoramidites **45**, **46** and **48** were used to make the substituted oligonucleotides shown in Table 3.1. The coupling efficiency of phosphoramidites **45**, **46** and **48** during automated DNA synthesis was determined. Oligonucleotide **70** incorporating the modified nucleoside **F** was tested for its ability to undergo post DNA assembly modification (Section 3.5). The thermodynamic stability of duplexes containing oligonucleotides **81-88** (Table 3.2) were evaluated (Section 3.4.3).

3.2.1 Determination of coupling efficiency of phosphoramidites 45, 46 and 48

Phosphoramidites **45**, **46** and **48** were dissolved in anhydrous acetonitrile to give 0.1 M solutions for synthesis of oligonucleotides **70-72** (Table 3.1) along with the unmodified complementary sequence **69**. All oligonucleotides were synthesised using standard solid phase synthesis methods on a 0.2 μmol scale.

Semi-quantitative synthesis yields were calculated for oligonucleotides **69-72** by 4,4'-dimethoxytrityl assays, shown in Table 3.1. The cleaved 4,4'-dimethoxytrityl cation ($DMTr^+$) during each synthesis cycle can be quantified using UV spectrophotometry.[65,138] (Chapter 6). The results of the 4,4'-dimethoxytrityl cation assay for oligonucleotides **69-72** are shown in Figure 3.4. The graphs show cycle number against 4,4'-dimethoxytrityl cation absorbance as a percentage of the second 4,4'-dimethoxytrityl cation reading. The second 4,4'-dimethoxytrityl cation reading is the preferred baseline reading as the first reading can a give misleading result due to 4,4'-dimethoxydetritylation of the support-bound nucleoside with time. As the solid support-bound base is 3'-linked (Figure 3.1), the 4,4'-dimethoxytrityl assays are carried out as synthesis progresses in the 3' to 5' direction. Overall synthetic yields for oligonucleotides **69-72**, as with all such calculations, were determined by calculating the ratio of the last and second 4,4'-dimethoxytrityl cation absorbance reading.

$$Overall\ yield = \frac{\text{Abs (DMTr cleaved after addition of the last nucleoside)}}{\text{Abs (DMTr cleaved after addition of the second nucleoside)}} \text{ x } 100$$

Coupling efficiencies for the novel phosphoramidites **45**, **46** and **48** were determined by calculating the ratio of the absorbance of the 4,4'-dimethoxytrityl cation cleaved after the addition of the modified nucleoside and the absorbance of the 4,4'-dimethoxytrityl cation calculated for the previous nucleoside addition.

$$Coupling\ yield = \frac{\text{Abs (DMTr cleaved after addition of the modified nucleoside)}}{\text{Abs (DMTr cleaved after addition of the previous nucleoside)}} \text{ x } 100$$

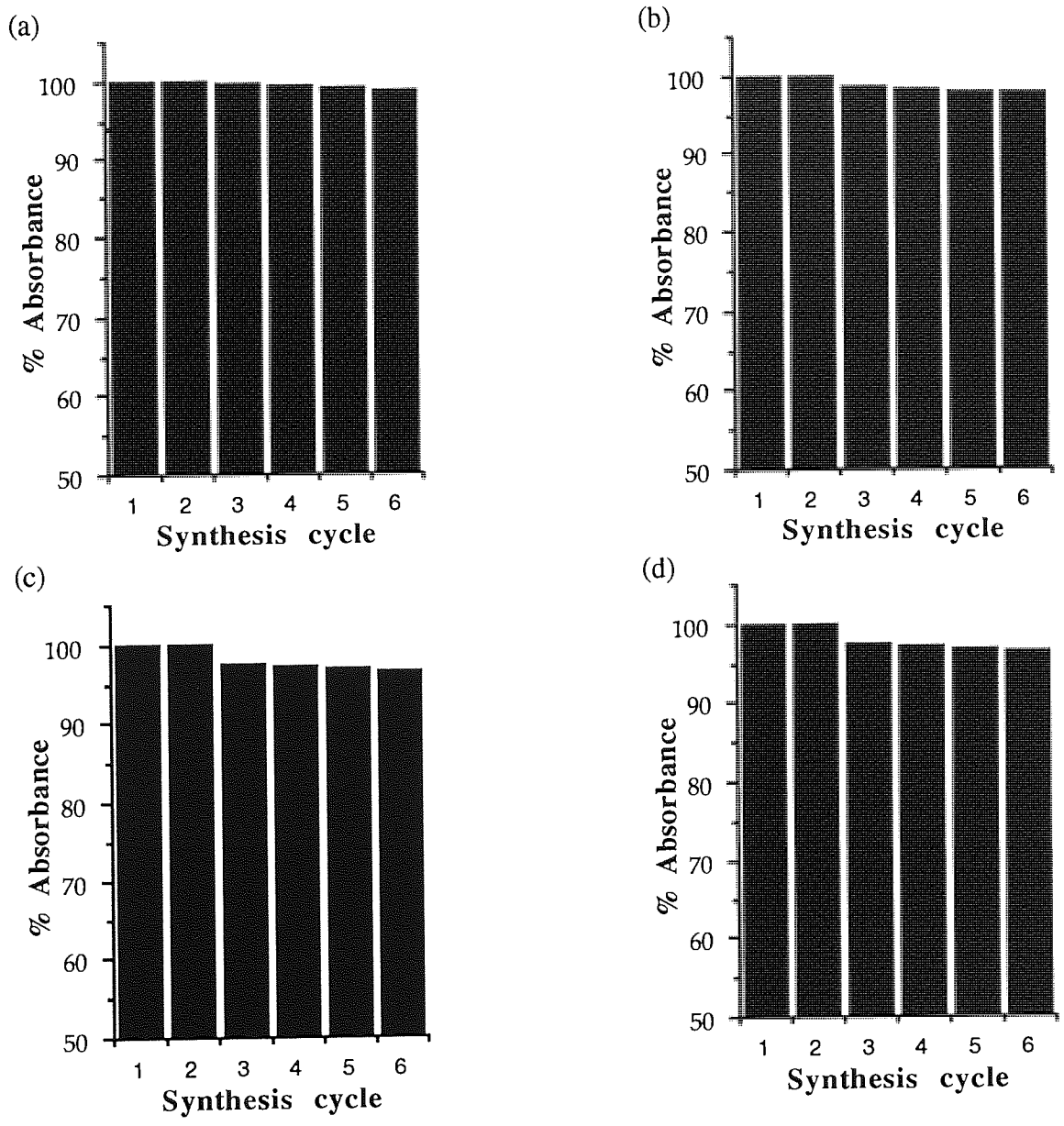


Figure 3.4 Percentage coupling yields for the synthesis of oligonucleotides (a) **69**; (b) **70**; (c) **71** and (d) **72**.

Oligonucleotide	Phosphoramidite	Sequence 5'-3'	Coupling yield	Synthesis yield
69	-	GCGCGC	-	98.9%
70	**45**	GCGFGC	98.9%	98.1%
71	**46**	GCGPGC	97.6%	96.7%
72	**48**	GCGMGC	97.8%	96.6%

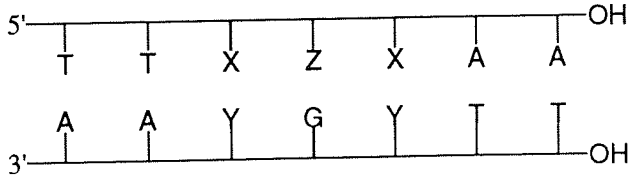
Table 3.1 Overall coupling yields for oligonucleotides **69-72**.

Standard operating conditions and wait times were used for the synthesis of oligonucleotides **69-72**. The coupling time for 3' to 5' addition of the modified nucleosides was 25 seconds. No attempt was made to improve coupling efficiencies by altering coupling time since acceptable yields were obtained.

Phosphoramidites **45**, **46** and **48** were shown to be compatible with automated DNA synthesis to make oligonucleotides **70-72** as judged by standard 4,4'-dimethoxytrityl cation analysis. Attempts to incorporate H-phosphonate contaminated phosphoramidite **40** into an oligonucleotide using standard DNA chemistry gave poor coupling yields in the order of 15% which rendered purification of the required oligonucleotide by HPLC impractical.

3.2.2 Synthesis of oligonucleotides for thermal analysis

A further series of oligonucleotides was synthesised to establish the effect of N4 aralkylation on thermodynamic stability. How the modified bases **P** and **M** affect thermodynamic stability of a particular oligonucleotide depends on its hydrogen bonding motif (sequence independent) and base-stacking potential (sequence dependent). The modified bases **P** and **M** were incorporated into non-self-complementary heptamers with either a C, G, A, T nucleosides above and below the plain of **P** and **M**, within the oligonucleotide (Figure 3.5).



Where **Z** = **P** or **M**

and X = A Y = T
X = T Y = A
X = G Y = C
X = C Y = G

Figure 3.5 Schematic representation of an oligonucleotide where Z indicates the position of the modified nucleoside (**P** or **M**) between bases X and Y.

Oligonucleotide	Sequence 3'-5'	Oligonucleotide	Sequence 3'-5'
73	TTGGGAA	**81**	TTGPGAA
74	TTTGTAA	**82**	TTTPTAA
75	TTAGAAA	**83**	TTAPAAA
76	TTCGCAA	**84**	TTCPCAA
77	TTGCGAA	**85**	TTGMGAA
78	TTTCTAA	**86**	TTTMTAA
79	TTACAAA	**87**	TTAMAAA
80	TTCCCAA	**88**	TTCMCAA

Table 3.2 Assignment of oligonucleotides for thermal analysis.

3.2.3 Purification and analysis of oligonucleotide purity

The oligonucleotides **69**, **71-88** were fully deprotected by suspending the derivatised CPG in concentrated ammonia (1 mL) and heated to 55°C for 18 hours (Figure 3.6). The CPG-bound oligonucleotide is liberated from the succinamide linker by base-catalysed hydrolysis (Figure 3.6 (a)). Deprotection of the 2-cyanoethyl protected phosphate groups is achieved by β-elimination (Figure 3.6 (b)) with concomitant removal of benzoyl protection from A and C and *iso*butyryl protection from G (Figure 3. 6 (c) and (d)). Oligonucleotides were dried by centrifugation under high vacuum and suspended in water (1 mL) for HPLC analysis and subsequent purification.

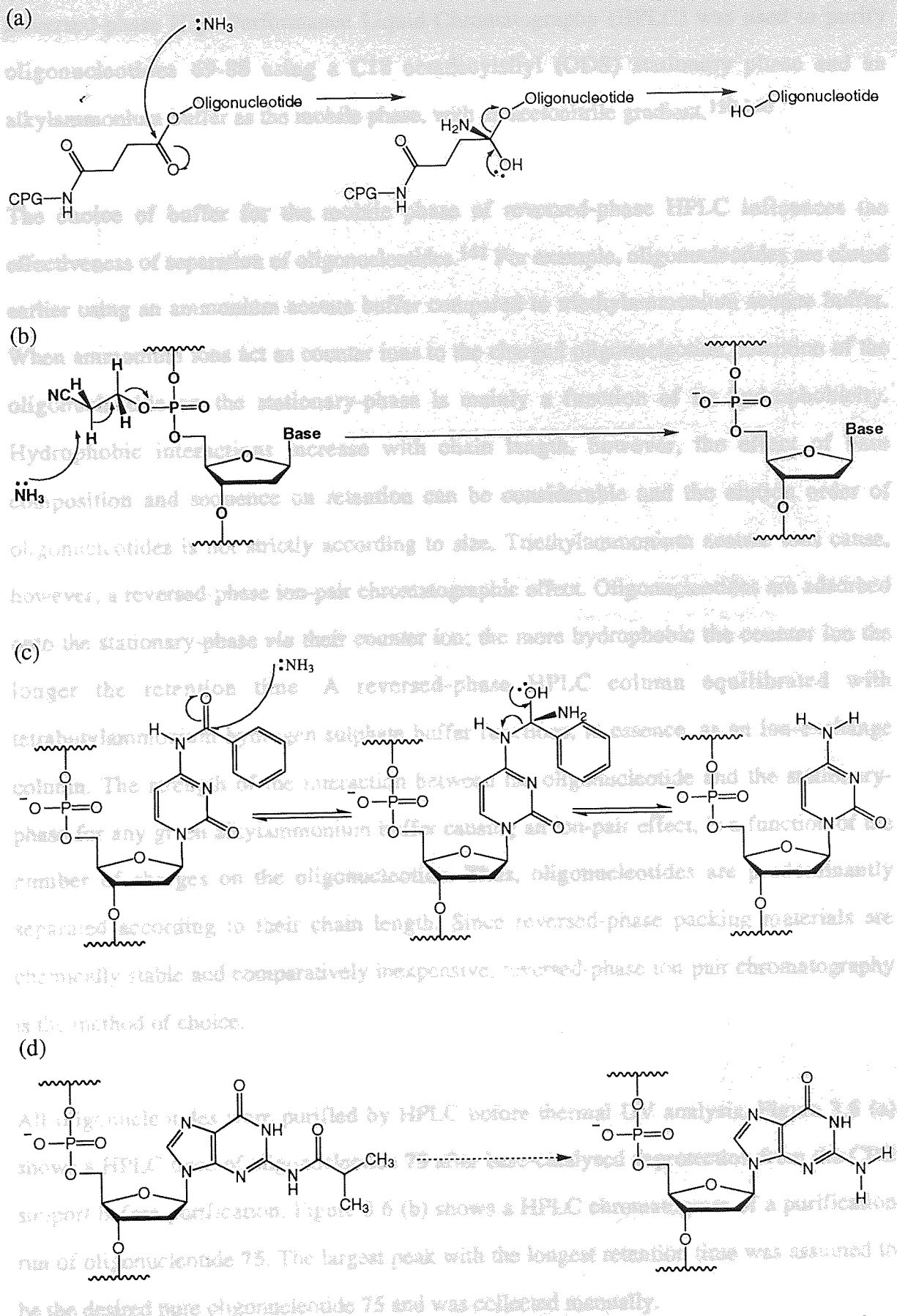


Figure 3.6 (a) Cleavage of the CPG-bound oligonucleotide; (b) β-elimination of cyanoethyl protected groups with concomitant deprotection of (c) benzoyl and (d) *iso*buytryl functions.

Reversed-phase High Performance Liquid Chromatography (HPLC) was used to purify oligonucleotides **69-88** using a C18 octadecylsilyl (ODS) stationary phase and an alkylammonium buffer as the mobile phase, with an acetonitrile gradient.[139,140]

The choice of buffer for the mobile phase of reversed-phase HPLC influences the effectiveness of separation of oligonucleotides.[141] For example, oligonucleotides are eluted earlier using an ammonium acetate buffer compared to triethylammonium acetate buffer. When ammonium ions act as counter ions to the charged oligonucleotide, retention of the oligonucleotide on the stationary-phase is mainly a function of its hydrophobicity. Hydrophobic interactions increase with chain length, however, the effect of base composition and sequence on retention can be considerable and the elution order of oligonucleotides is not strictly according to size. Triethylammonium acetate ions cause, however, a reversed-phase ion-pair chromatographic effect. Oligonucleotides are adsorbed onto the stationary-phase *via* their counter ion; the more hydrophobic the counter ion the longer the retention time. A reversed-phase HPLC column equilibrated with tetrabutylammonium hydrogen sulphate buffer functions, in essence, as an ion-exchange column. The strength of the interaction between the oligonucleotide and the stationary-phase for any given alkylammonium buffer causing an ion-pair effect, is a function of the number of charges on the oligonucleotide. Thus, oligonucleotides are predominantly separated according to their chain length. Since reversed-phase packing materials are chemically stable and comparatively inexpensive, reversed-phase ion-pair chromatography is the method of choice.

All oligonucleotides were purified by HPLC before thermal UV analysis. Figure 3.6 (a) shows a HPLC trace of oligonucleotide **75** after base-catalysed deprotection from the CPG support before purification. Figure 3.6 (b) shows a HPLC chromatogram of a purification run of oligonucleotide **75**. The largest peak with the longest retention time was assumed to be the desired pure oligonucleotide **75** and was collected manually.

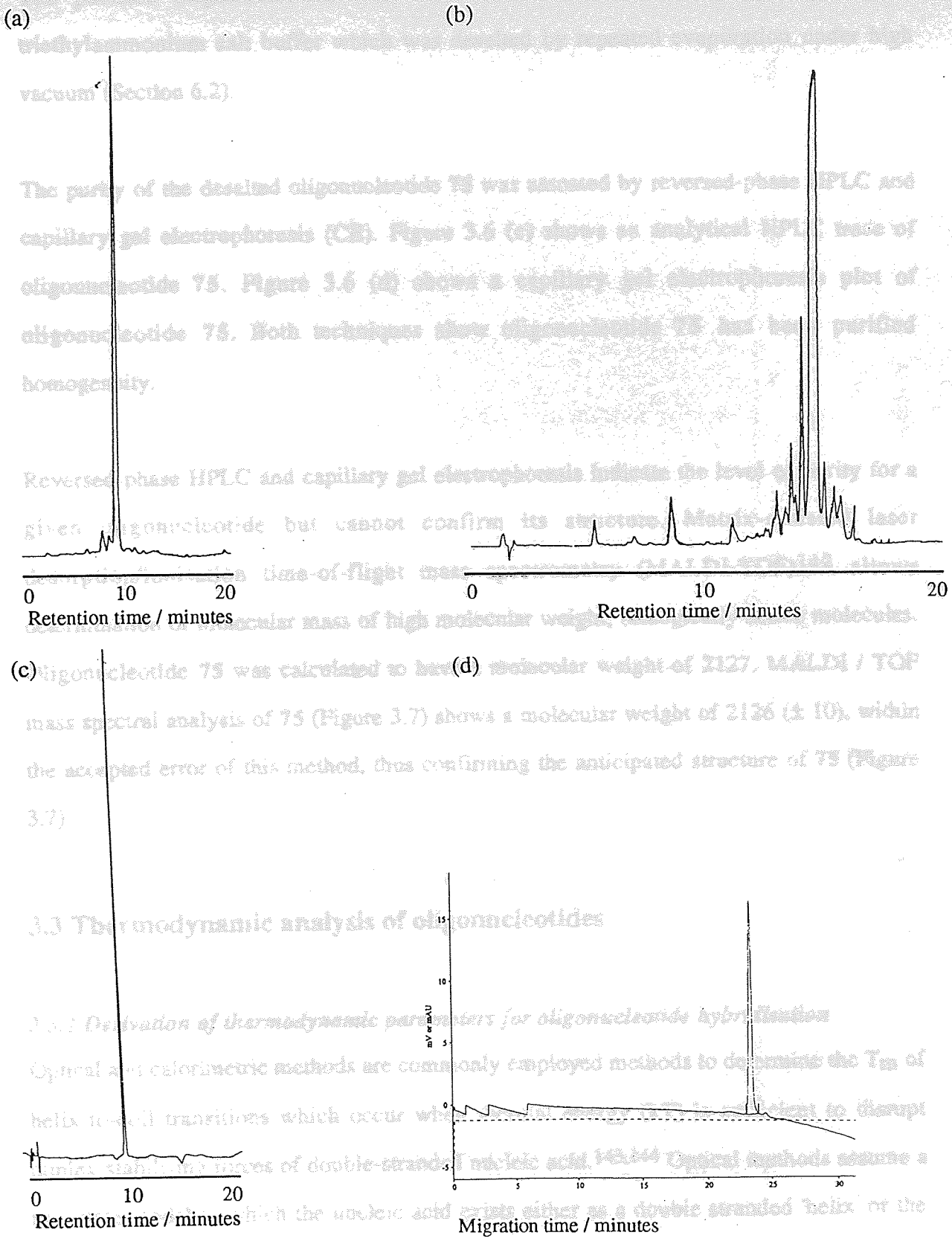


Figure 3.6 HPLC profile of (a) crude oligonucleotide **75** after deprotection and cleavage from the solid support; (b) purification run of **75**; (c) **75** after purification and (d) CE analysis of purified **75**.

The purified oligonucleotide was collected as a solute in an aqueous 0.1 M triethylammonium salt buffer which was desalted by repeated evaporation under high vacuum (Section 6.2).

The purity of the desalted oligonucleotide **75** was assessed by reversed-phase HPLC and capillary gel electrophoresis (CE). Figure 3.6 (c) shows an analytical HPLC trace of oligonucleotide **75**. Figure 3.6 (d) shows a capillary gel electrophoresis plot of oligonucleotide **75**. Both techniques show oligonucleotide **75** has been purified homogeneity.

Reversed-phase HPLC and capillary gel electrophoresis indicate the level of purity for a given oligonucleotide but cannot confirm its structure. Matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry (MALDI-TOF)[142] allows determination of molecular mass of high molecular weight, biologically active molecules. Oligonucleotide **75** was calculated to have a molecular weight of 2127. MALDI / TOF mass spectral analysis of **75** (Figure 3.7) shows a molecular weight of 2126 (± 10), within the accepted error of this method, thus confirming the anticipated structure of **75** (Figure 3.7).

3.3 Thermodynamic analysis of oligonucleotides

3.3.1 Derivation of thermodynamic parameters for oligonucleotide hybridisation

Optical and calorimetric methods are commonly employed methods to determine the T_m of helix-to-coil transitions which occur when thermal energy (kT) is sufficient to disrupt duplex-stabilising forces of double-stranded nucleic acid.[143,144] Optical methods assume a two-state model in which the nucleic acid exists either as a double stranded 'helix' or the single stranded 'coil', and that no intermediate states exist. Deviation from the two-state model occurs due to internal and terminal melting of the nucleic acid duplexes, which reduces the reliability of the results (Figure 3.8).

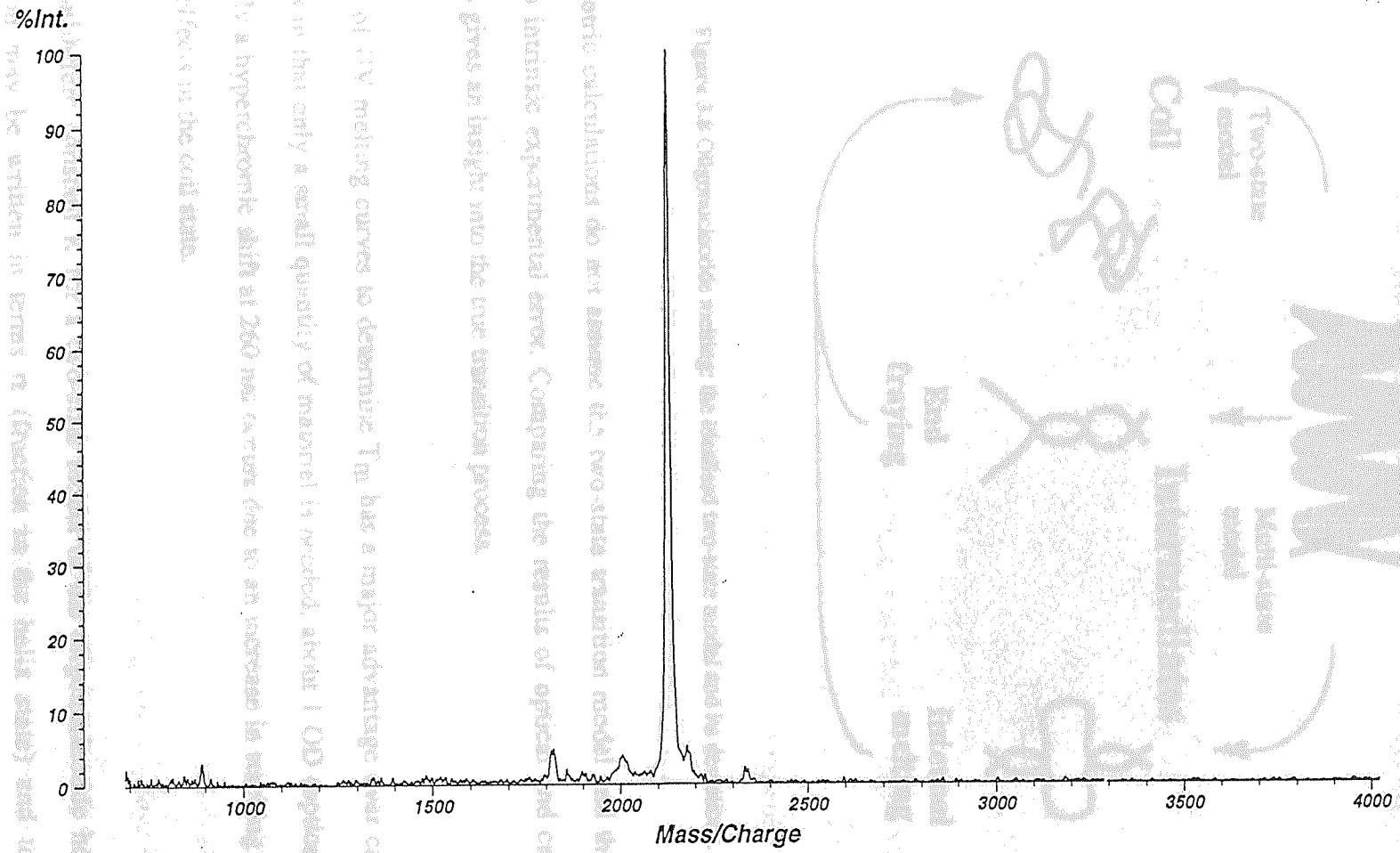


Figure 3.7 MALDI-TOF mass spectral analysis of oligonucleotide **75**.

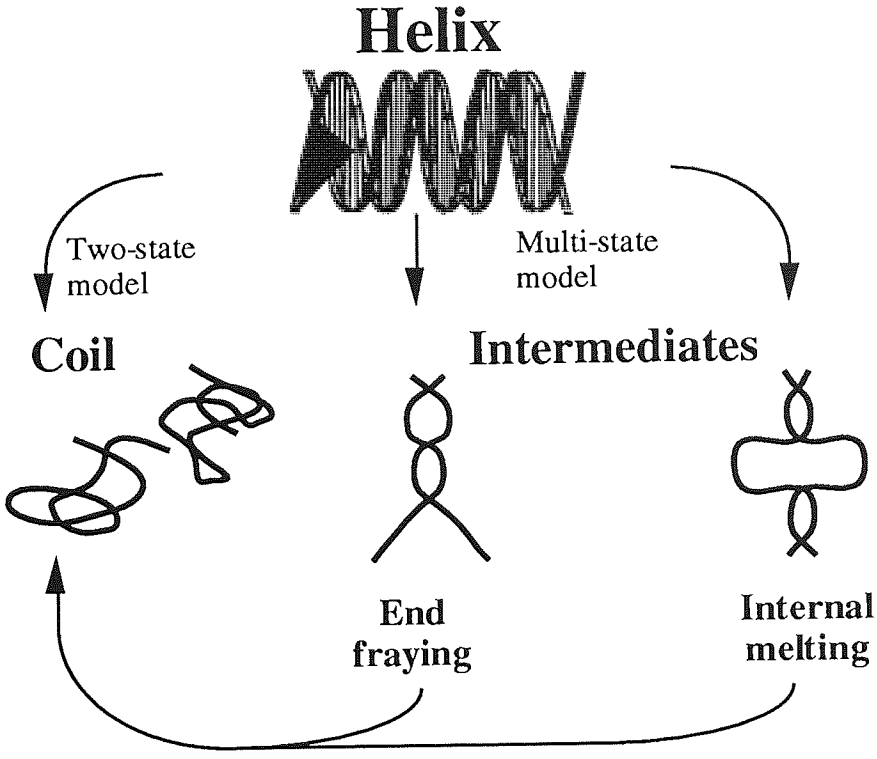


Figure 3.8 Oligonucleotide melting: the idealised two-state model and its deviations.

Calorimetric calculations do not assume the two-state transition model and thus are not liable to intrinsic experimental error. Comparing the results of optical and calorimetric methods gives an insight into the true transition process.

The use of UV melting curves to determine T_m has a major advantage over calorimetric methods in that only a small quantity of material is needed, about 1 OD (optical density). Generally a hyperchromic shift at 260 nm occurs due to an increase in stacking dependant optical effects in the coil state.

The equilibrium constant K for a two-state model of the oligonucleotide helix-to-coil transition may be written in terms α (fraction in the helix state) and total strand concentration C_t. The equilibrium constant for a self-complementary strand such as CGCGCG is given by Equation 6. The equilibrium constant of non-identical

complementary strands is given by Equation 7 while duplexes formed intramolecularly, for example, GGGAAAACCC are given by Equation 8.

$$K = \frac{\alpha / 2}{(1-\alpha)^2 C_t} \qquad \text{[Eqn. 6]}$$

$$K = \frac{2\alpha}{(1-\alpha)^2 C_t} \qquad \text{[Eqn. 7]}$$

$$K = \frac{\alpha}{(1-\alpha)} \qquad \text{[Eqn. 8]}$$

The T_m value occurs when $\alpha = 0.5$. Substitution of Equations 6 and 7 into the Van't Hoff's isochore Equation 9, the integrated form of Equation 10, yields Equation 11 which can be used to determine the Van't Hoff's enthalpy for intermolecular duplex formation.

$$\frac{d \ln K_p}{dT} = \frac{\Delta H^\circ}{RT^2} \qquad \text{[Eqn. 9]}$$

$$\ln \frac{K_2}{K_1} = -\frac{\Delta H^\circ}{R}\left(\frac{1}{T_2} - \frac{1}{T_1}\right) \qquad \text{[Eqn. 10]}$$

$$\Delta H_{vH} = 6RT^2(\delta\alpha / \delta T) \qquad \text{[Eqn. 11]}$$

The gradient at the T_m inflection point gives a value of $\delta\alpha/\delta T$ on a melting curve plot of 1-α versus T. Although the use of Equation 11 to calculate ΔH_{vH} is convenient, there are problems associated with it. The assumption made is that melting is a two-stage process: all or nothing. While this is probably true for short oligonucleotides, larger ones may undergo localised melting and end fraying. Progression through a number of intermediate stages leads to a broadened plot of 1-α versus T and a lower ΔH_{vH}.

This method using Van't Hoff's equation assumes a temperature-independent enthalpy which is not necessarily the case. Another drawback is that the baseline is not flat before and after the graphical turning points. Choice of baseline can affect the result by 20%. This problem can be overcome by plotting the differential ($\delta\alpha/\delta T$) against T. The half-width of this curve is inversely proportional to the Van't Hoff transition enthalpy at the half height. This method overcomes the problems of baseline distortions but still assumes a two state transition and a temperature-independent enthalpy.

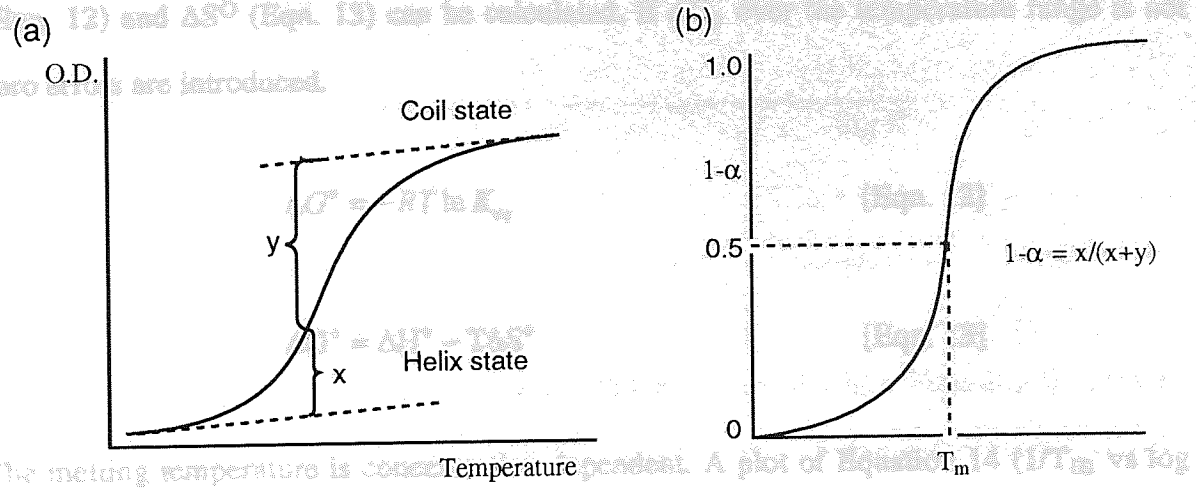


Figure 3.9 (a) Sigmoidal UV melting curve and (b) its 1-α plot.

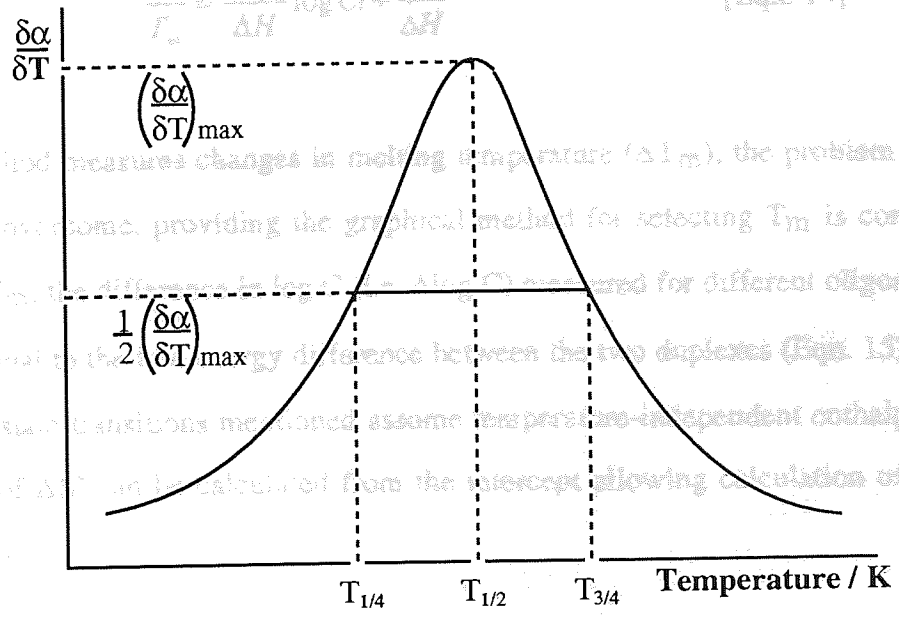


Figure 3.10 Derivative UV melting plot.

The equilibrium constant is dependant on temperature. If ΔH_{vH} is constant over a range then a plot of ln |K| against 1/T is linear. A value for ΔH_{vH} can be obtained by multiplying the gradient by -R. Values of ln |K| are taken for plots of 1-α versus T using Equations 6 and 7. As stated already, substituting $\alpha = 0.5$ into the appropriate rate equations (Eqn. 6, 7, or 8) will give the equilibrium constant corresponding to the T_m. Using the appropriate substitution in the integrated Van't Hoff's isochore (Eqn. 10), the standard equilibrium constant at 25°C can be extrapolated. Using standard thermodynamic relationships, ΔG^O (Eqn. 12) and ΔS^O (Eqn. 13) can be calculated. If ΔC_p over the temperature range is not zero errors are introduced.

$$\Delta G^o = -RT \ln K_{eq} \qquad \text{[Eqn. 12]}$$

$$\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ \qquad \text{[Eqn. 13]}$$

The melting temperature is concentration dependent. A plot of Equation 14 ($1/T_m$ vs log C_t) is linear, the intercept giving $\Delta S/\Delta H$ and the gradient giving $2.3R/\Delta H$.

$$\frac{1}{T_m} = \frac{2.3R}{\Delta H} \log C_t + \frac{\Delta S}{\Delta H} \qquad \text{[Eqn. 14]}$$

As this method measures changes in melting temperature (ΔT_m), the problem of variable baseline is overcome, providing the graphical method for selecting T_m is constant. At a particular T_m, the difference in log C (i.e. Δlog C) measured for different oligonucleotides, is proportional to the free energy difference between the two duplexes (Eqn. 15). All of the above two-state transitions mentioned assume temperature-independent enthalpy changes. The value of ΔS° can be calculated from the intercept allowing calculation of ΔG° using Equation 10.

$$\Delta G_i^\circ - \Delta G_j^\circ = 2.3RT\left(\log C_T^i + \log C_T^j\right) \qquad \text{[Eqn. 15]}$$

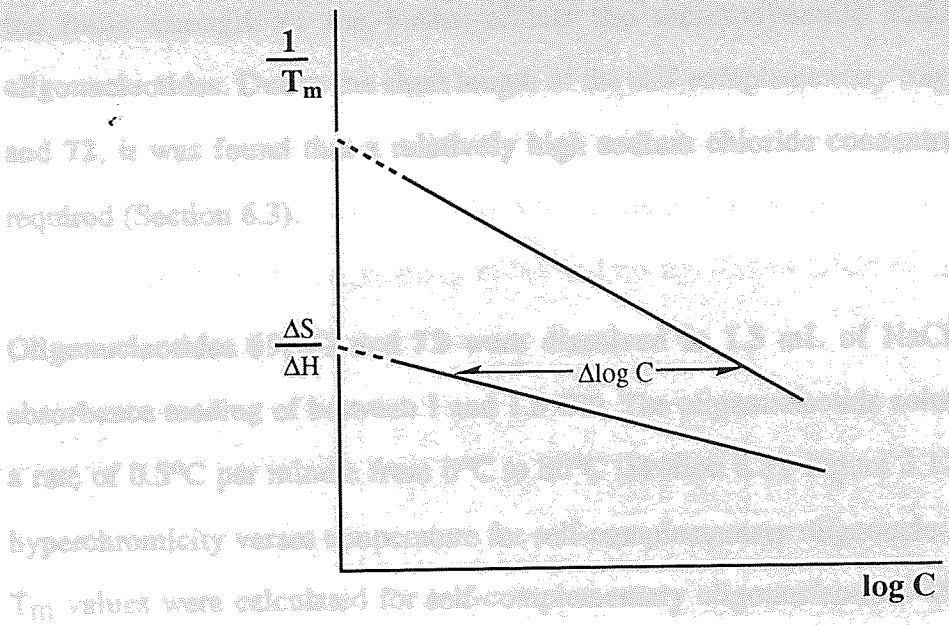


Figure 3.11 Δlog C is proportional to the free energy difference between the two duplexes.

For intramolecular duplex formation, substitution of $\alpha = 0.5$ into Equation 9 gives an equilibrium constant of one, consequently ΔG^{O} is then zero allowing determination of ΔS^{O} using Equation 17.

$$\Delta S^{\circ} = \frac{\Delta H^{\circ}}{T_m} \qquad \text{[Eqn. 16]}$$

Substituting Equation 16 into Equation 13 gives allows determination of Gibb's free energy at any temperature (Eqn. 17).

$$\Delta G^{\circ} = \Delta H\left(1 - \frac{T}{T_m}\right) \qquad \text{[Eqn. 17]}$$

3.3.2 Thermodynamic analysis of self-complementary oligonucleotides 69, 71 and 72

In this work, the self-complementary oligonucleotide sequences **69**, **71** and **72** were subjected to UV thermal analysis to determine their T_m values (Figure 3.12). The choice of buffer for oligonucleotide thermodynamic analysis is of importance. As stated (Chapter 1),

the ionic strength of the buffer affects the thermodynamic stability of hybridised oligonucleotides. Due to the short length of the self-complementary oligonucleotides **69**, **71** and **72**, it was found that a relatively high sodium chloride concentration of 1.0 M was required (Section 6.3).

Oligonucleotides **69**, **71** and **72** were dissolved in 1.3 mL of NaCl buffer to give an absorbance reading of between 1 and 1.5 OD. The oligonucleotide solutions were heated at a rate of 0.5°C per minute from 0°C to 80°C (Section 6.3). Figure 3.12 shows a plot of % hyperchromicity verses temperature for self-complementary oligonucleotides **69**, **71** and **72**. T_m values were calculated for self-complementary oligonucleotides **69** and **71** by plotting their 1-α curves and taking a reading at $\alpha = 0.5$ (Table 3.3). Observation of the thermal analysis curve for self-complementary oligonucleotide **72** showed introduction of nucleoside **M** had such a destabilising effect on duplex stability that no reliable T_m value could be measured.

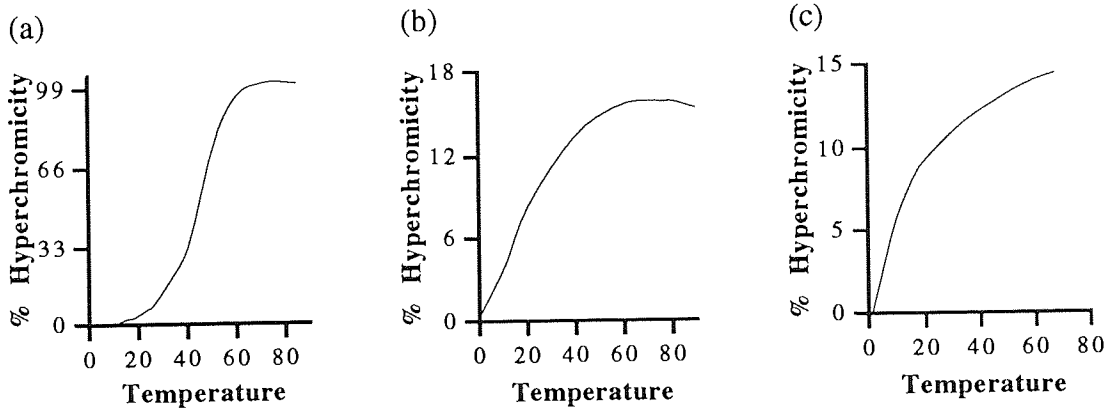


Figure 3.12 UV melting profiles of self-complementary oligonucleotides (a) **69**; (b) **70** and (c) **71**.

Oligonucleotide	**69**	**71**	**72**
OD	1.5	1.4	1.2
T_m/°C ± 0.2	54	25	-

Table 3.3 Tm values of self-complementary oligonucleotides **69**, **71** and **72**.

Concentration of oligonucleotides **69**, **71** and **72** were determined by taking the sum of the individual extinction coefficients (ε). The coefficient ε of nucleoside **P** was determined using the Beer-Lambert law at different wavelengths (λ) and temperatures (Table 3.4). The extiction coefficient of nucleoside **M** were calculated in the same manner as for nucleoside **P**. The addition of a methyl group at N4 had no significant effect on the value of ε, so the extinction coefficient of **M** was taken to be that of **P.**

λ/nm	ε at 0 °C	ε at 25 °C	ε at 50 °C	ε at 75 °C	ε at 100 °C
254.4	9700 ± 190	10000 ± 200	10300 ± 200	10000 ± 200	9800 ± 200
260.0	10300 ± 200	10000 ± 200	10800 ± 200	10800 ± 220	10500 ± 210
270.0	13700 ± 270	12600 ± 250	13300 ± 260	13200 ± 264	12800 ± 260

Table 3.4 Extinction coefficients for nucleoside **P** at different wavelengths (λ) and temperatures.

The T_m vales of self-complementary oligonucleotides **69** and **71** were calculated a varying oligonucleotide concentrations (Table 3.5).

Oligonucleotide	**Concentration x10^6**	**T_m/°C ± 0.2**
69	19.9	54.1
	14.7	52.0
	12.3	50.3
	8.91	48.5
	7.10	47.0
71	8.77	25.2
	6.82	22.7
	5.62	21.2
	4.78	19.8
	3.34	21.2

Table 3.5 T_m as a function of concentration for the self-complementary oligonucleotides **69** and **71.**

Thermodynamic parameters ΔH, ΔS and ΔG for oligonucleotides **69** and **71** were calculated using the integrated Van't Hoff's isochore (Equation 14). Figure 3.13 shows a plot of 1/Tm against log concentration for nucleosides **69** and **71**.

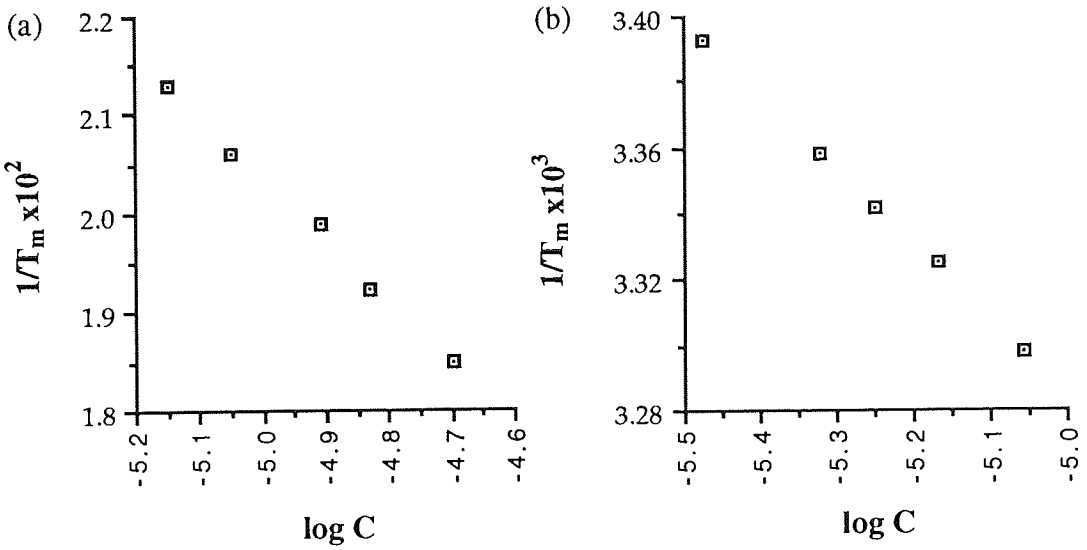


Figure 3.13 Plot of temperature verses concentration in accordance with the integrated Van't Hoff's isochore (Equation 14) for (a) **69** and (b) **71**.

The change in enthalpy (ΔH) of duplex melting is a function of the gradient of Equation 14, assuming ΔC_p over the temperature range is zero. Likewise, the change in entropy (ΔH) may be calculated from the point of intersection with the y axis. Thermodynamic parameters for the self-complementary oligonucleotides **69** and **71** are given in Table 3.6 and are displayed graphically in Figure 3.14.

Oligonucleotide	ΔG/KCal.mol^{-1}	ΔH/KCal.mol^{-1}	TΔS/KCal.mol^{-1}
69	-9.8 ± 7	31.8 ± 1	22.5 ± 6
71	-8.2 ± 5	20.0 ± 0.5	13.4 ± 4.5

Table 3.6 Thermodynamic parameters for the self-complementary oligonucleotides **69** and **71**.

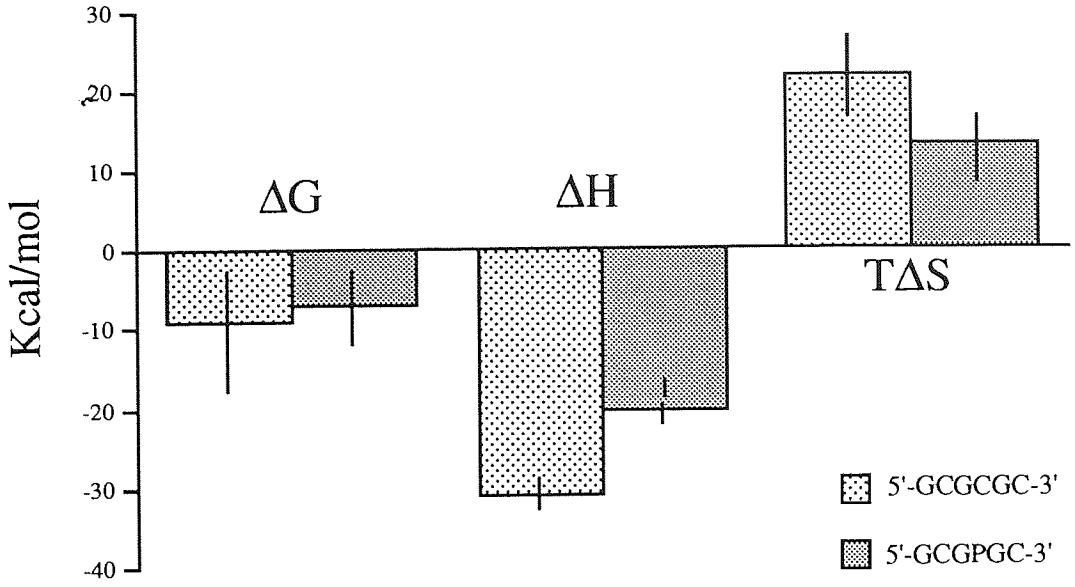


Figure 3.14 Thermodynamic parameters for the self-complementary oligonucleotides **69** and **71**.

Observation of the UV melting-curves and consideration of the thermodynamic parameters for the self-complementary oligonucleotides **69** and **71** shows that the inclusion of nucleoside **P** within the sequence of **71** had a thermodynamically destabilising effect with respect to helical stability. However, it is noted that while the enthalpic component of duplex stability is less favourable for the self-complementary oligonucleotide **71** than **69**, the entropic component was more favourable (i.e. less disruptive) to the overall duplex stability of **71**. Gibb's free energy for duplexation (ΔG) is a result of large but opposing enthalpy and entropy terms; in the case of **71**, the benefits to duplex stability of an improved entropy contribution were outweighed by the loss of enthalpic stabilisation.

3.3.3 Thermodynamic analysis of non self-complementary oligonucleotides 81-88

In an attempt to identify the effect of different bases above and below the plane of nucleoside **P**, oligonucleotides **77-88** were combined in a 1:1 ratio with their self-complementary sequences **73-76** (Section 6.3). Figure 3.15 show the UV melting curves of modified oligonucleotides **77-88** with their complementary sequences in comparison with their 'parent' unmodified sequences (Table 3.7).

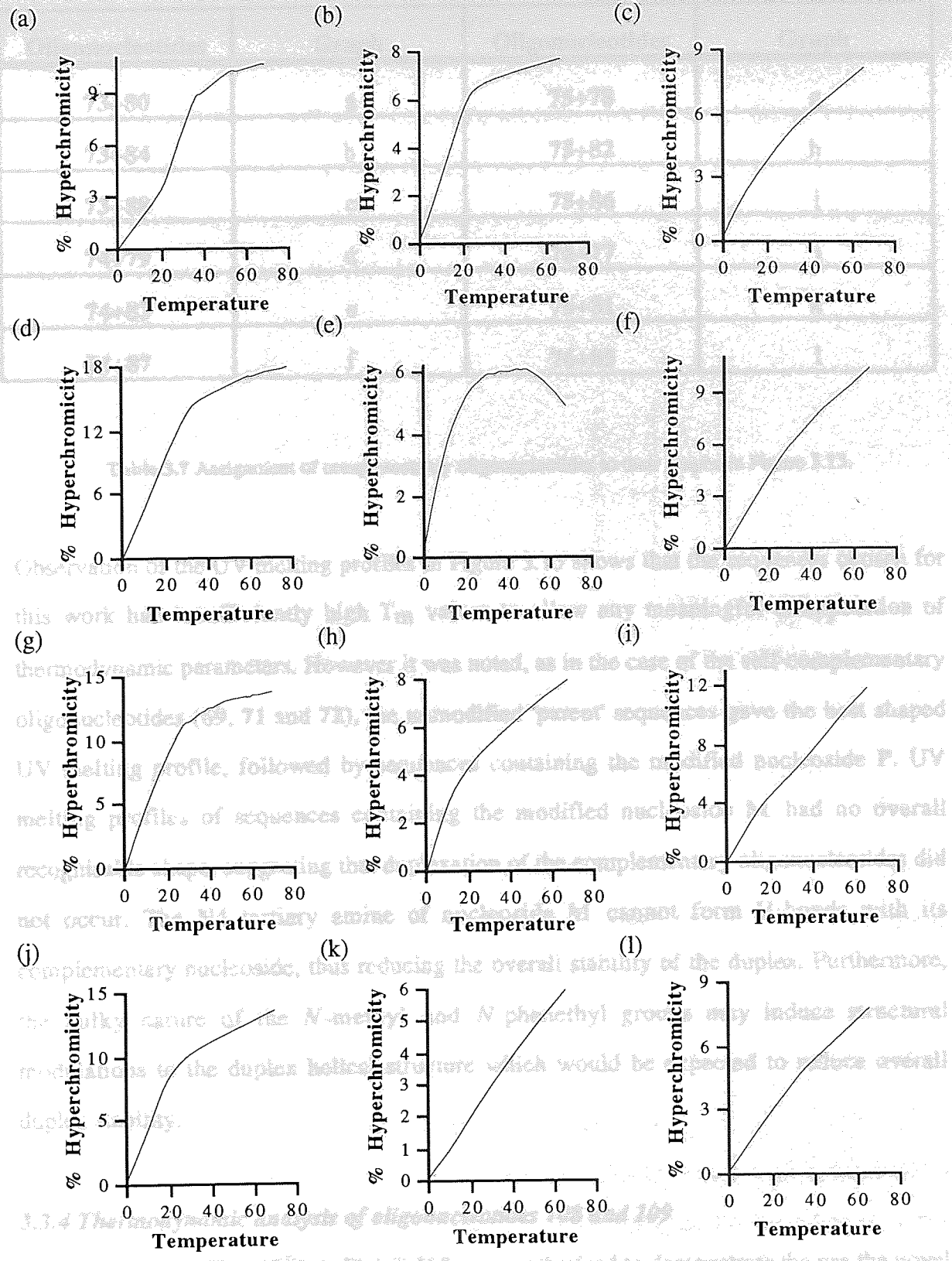


Figure 3.15 UV melting-curves for oligonucleotides **77-88** and their complementary sequences.

Oligonucleotides	Graph	Oligonucleotides	Graph
73+80	a	**75+78**	g
73+84	b	**75+82**	h
73+88	c	**75+86**	i
74+79	d	**76+77**	j
74+83	e	**76+81**	k
74+87	f	**76+85**	l

Table 3.7 Assignment of complementary oligonucleotides to their graphs in Figure 3.15.

Observation of the UV melting profiles in Figure 3.15 shows that the sequences chosen for this work had insufficiently high T_m values to allow any meaningful extrapolation of thermodynamic parameters. However it was noted, as in the case of the self-complementary oligonucleotides (**69**, **71** and **72**), the unmodified 'parent' sequences gave the best shaped UV melting profile, followed by sequences containing the modified nucleoside **P**. UV melting profiles of sequences containing the modified nucleoside **M** had no overall recognisable shape, suggesting that duplexation of the complementary oligonucleotides did not occur. The N4 tertiary amine of nucleoside **M** cannot form H-bonds with its complementary nucleoside, thus reducing the overall stability of the duplex. Furthermore, the bulky nature of the *N*-methyl and *N*-phenethyl groups may induce structural modulations to the duplex helical structure which would be expected to reduce overall duplex stability.

3.3.4 Thermodynamic analysis of oligonucleotides 108 and 109

Oligonucleotide $d(Tp)_3(C^{Bz}p)_2(Tp)_2T$ **108** was synthesised to demonstrate the use the novel Controlled Glass Pore support bound silyl linker **106** (Chapter 5). Its structure was confirmed by MALDI-TOF mass spectrometric analysis. Oligonucleotide **109**, the

complementary sequence to **108**, was synthesised using a standard succinamide linked solid support.

Oligonucleotides **107** and **108** were mixed in a 1:1 ratio with **109** in 1.5 mL of 1.0 M NaCl buffer (Section 6.3) to give an overall OD of 1.5 for each solution. UV thermal analysis was carried out following the procedure outlined in section 3.4.2. Figure 3.16 shows the melting profiles for oligonucleotides (a) **107+109** and (b) **108+109**.

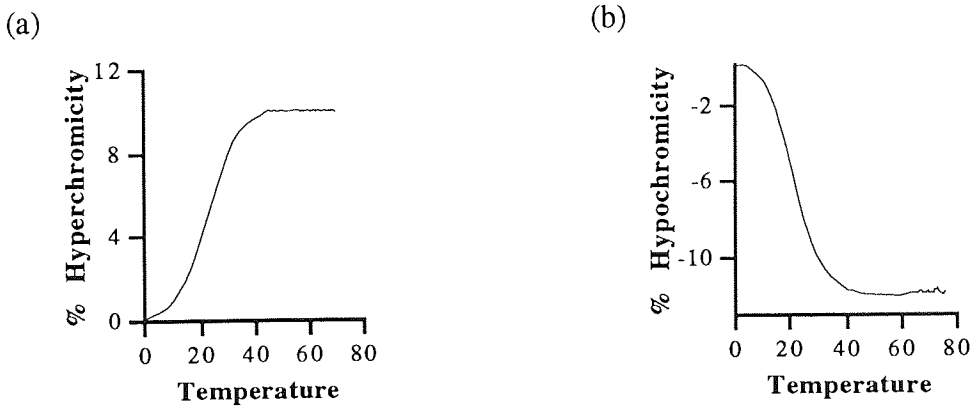


Figure 3.16 UV melting profiles for oligonucleotides (a) **107+109** and (b) **108+109**.

Although oligonucleotides **107+109** appeared to have annealed to form a helical duplex by the sigmoidal shape of the hypochromic UV melting curve, the helix was destabilised having a T_m of 22°C compared to the 'parent', unsubstituted duplex **107+108**, which had a T_m of 32°C.

3.4 Post-synthetic substitution of oligonucleotide 70

The ability of nucleoside **F** to undergo nucleophilic attack by amine nucleophiles whist incorporated in an oligonucleotide was examined. Oligonucleotide **70**, bound to the CPG solid support *via* a base-sensitive succinamide linker, was treated with saturated aqueous ammonia solution for 20 hours at 60°C. It was expected that treatment with conc. NH_3 (aq.) would serve to cleave the oligonucleotide from the solid support and deprotect the heterocyclic bases and cyanoethyl-protected phosphorous centres. Ammonia was also

expected to act as a nucleophile to displace the pentafluorophenoxy group at C4 of nucleoside **F** (Figure 3.17).

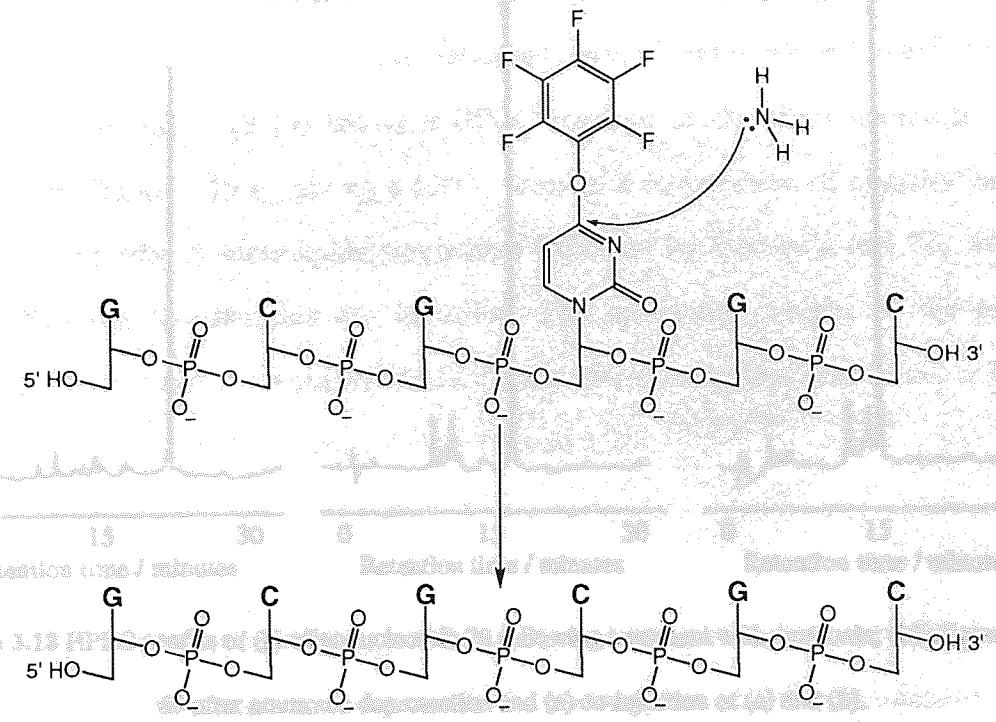


Figure 3.17 Post-synthetic modification of oligonucleotide **70**.

After treatment with conc. NH_3 (aq.) the solid CPG was filtered from the ammonia solution using a 0.4 μm pore syringe filter. The supernatant was concentrated under vacuum and suspended in HPLC grade water (1 mL) and subjected to reversed-phase HPLC analysis. Figure 3.18 (a) shows a HPLC trace of oligonucleotide **70** after treatment with ammonia. Figure 3.18 (b) shows a HPLC trace of crude oligonucleotide **69** after deprotection. Figure 3.18 (c) shows a HPLC trace of a co-injection of oligonucleotide **69** and **70** after treatment with ammonia, which co-elute, indicating the samples are identical.

The susceptibility of nucleoside **F** to displacement by substituted amines was examined. Oligonucleotide **70** bound to the CPG solid support was suspended in *N*-methylphenethylamine (0.4 mL) and heated to 60°C for 20 hours. The solid support was isolated using a 0.4 μm pore syringe filter to which 1 mL of saturated aqueous ammonia

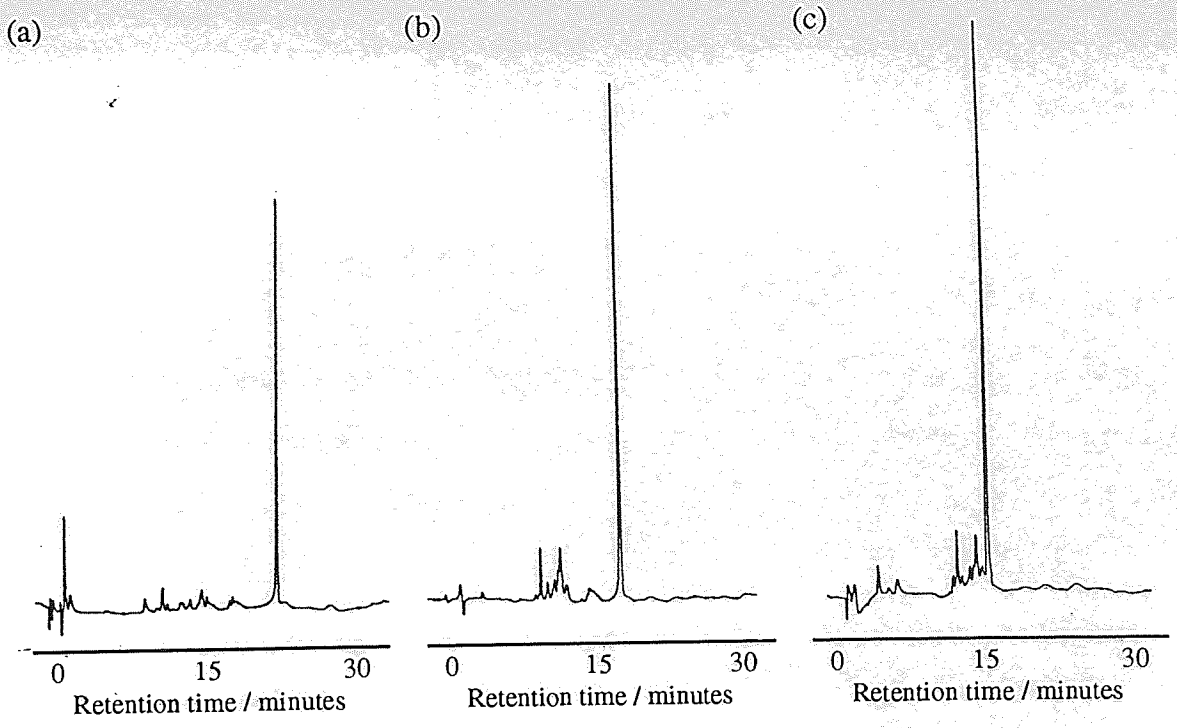


Figure 3.18 HPLC profile of (a) oligonucleotide **70** following treatment with ammonia; (b) oligonucleotide **69** after ammonia deprotection and (c) co-injection of (a) and (b).

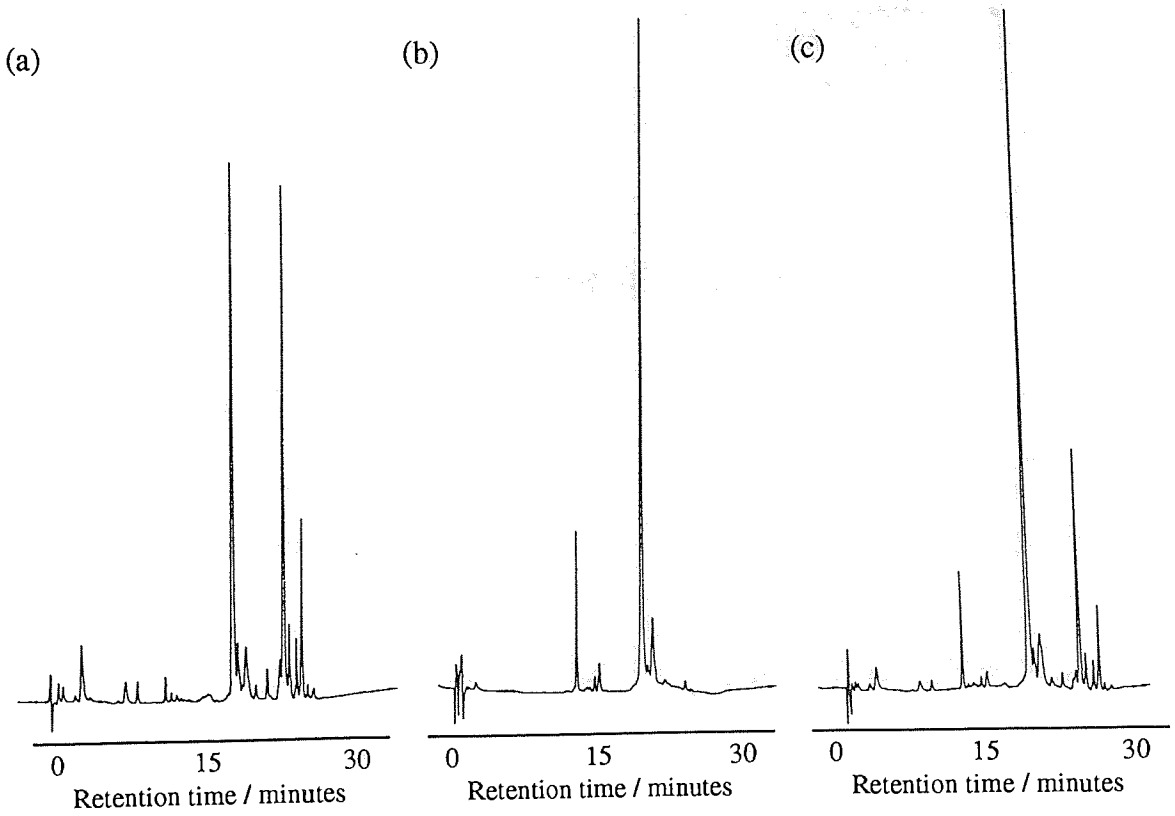


Figure 3.19 HPLC profile of (a) oligonucleotide **70** following treatment with *N*-methyl-*N*-phenethylamine and ammonia; (b) oligonucleotide **72** following ammonia deprotection and (c) co-injection of (a) and (b).

solution was added which was heated at 55°C for 18 hours. The solid support was separated from the supernatant which was concentrated under vacuum and suspended in 1 mL of HPLC grade water and subjected to reversed-phase HPLC analysis. Figure 3.19 (a) shows a HPLC trace of oligonucleotide **70** after treatment with *N*-methylphenethylamine followed by ammonia. Figure 3.19 (b) shows a HPLC trace of crude oligonucleotide **72** after deprotection. Figure 3.19 (c) shows a HPLC trace of a co-injection of oligonucleotide **70** after treatment with *N*-methylphenethylamine followed by ammonia and **72,** which co-elute, indicating the samples are identical. The molecular weight of the modified oligonucleotide **72** was analysed by MALDI mass spectrometry and was found to be 1910, identical to the calculated molecular weight (Figure 3.20).

3.5 Conclusion

Phosphoramidites **45**, **46** and **48** were fully compatible with solid phase synthesis with no additional protecting group required to mask the NH at C4 in **46**. All oligonucleotides were purified to homogeneity by reversed-phase HPLC; some of which were subjected to MALDI-TOF analysis to give unambiguous proof of their structural integrity.

Due to their short length, the hybridisation of only a modest number of oligonucleotides synthesised could be studied successfully using variable temperature UV spectrophotometry. Aralkyl substitutions invariably resulted in thermodynamic destabilisation of duplexes and were therefore inappropriate for antisense purposes.

Post-synthetic substitution of oligonucleotide **70** to give an N4-substituted cytosine was achieved. This post-synthetic substitution method could be developed for the incorporation other, beneficial substituents into oligonucleotides for biological and antisense applications.

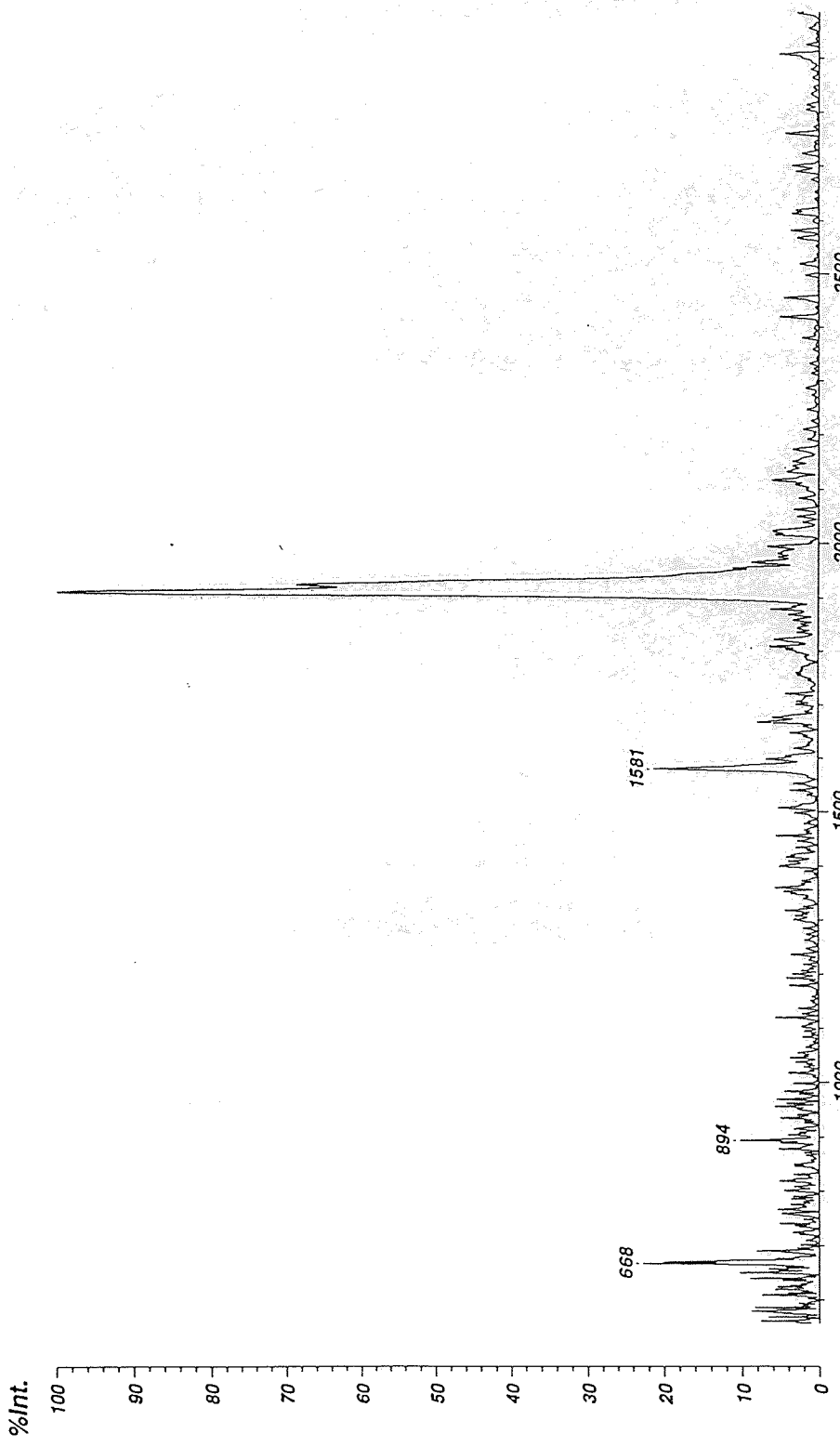


Figure 3.20 MALDI-TOF mass spectral analysis of oligonucleotide **72**.

Chapter 4

Direct Synthesis, Structure and Anti-HIV Activity of Modified Pyrimidine Nucleosides

4.1 Project Rationale

Pyrimidine nucleosides with displaceable groups at C4 are important intermediates for the synthesis of anti-viral agents,[145-148] anti-cancer compounds[149], oligonucleotides[150] and radiolabelled adducts.[151] To date, successful formation of C4 azolyl derivatives of 2'-deoxy nucleosides has required time consuming protection/deprotection strategies for the reactive hydroxyl groups on the sugar ring. Protecting groups have included acetyl, benzoyl, di-*tert*-butylmethylsilyl, 1,1,3,3-tetra*iso*propyl-1,3-disiloxanediyl, trityl, 4,4'-dimethoxytrityl and phosphoramidite.[119-122,125-127,152] Ribose nucleosides also require protection of the 2'-OH where the *tert*-butyldimethylsilyl[153] and tetrahydropyranyl groups have been successfully used.[154]

Although fairly common in the literature, 1,2,4-triazole (TZL) pyrimidine nucleoside intermediates have remained poorly characterised. This work was aimed at optimising the conditions necessary to synthesise both C4 pentafluorophenoxy and 1,2,4-triazolyl pyrimidines and at examining their anti-HIV activities.

4.2 Synthesis of C4 pentafluorophenoxy and 1,2,4-triazolyl pyrimidines

The methodology of Sung (1981)[120] for synthesis of nucleoside **2** requires acetyl protection of reactive hydroxyls and was applied, in this work, to 2'-deoxythymine, 2'-deoxyuridine and AZT in order to introduce the pentafluorophenoxy and 1,2,4-triazolyl groups at C4 (Table 4.1). All acetyl-protected pyrimidine intermediates have previously been described (Chapter 2) except for C5'-acetyl-protected AZT derivative **89** which was synthesised using standard procedures. Towards the synthesis of phosphoramidite **46**, two

one-pot methods were developed to introduce the displaceable pentafluorophenoxy group at C4 of 2'-deoxyuridine (Section 2.4.4 and 2.4.5). These methods were applied to 2'-deoxythymidine, AZT and to uridine in order to introduce the 1,2,4-triazolyl and pentafluorophenoxy groups at C4 (Table 4.1).

Figure 4.1 and Table 4.1 summarise the synthetic work undertaken. Purification of the C4-modified nucleosides was achieved by flash chromatography. All C4-modified pyrimidine nucleosides were tested as potential anti-HIV agents. Nucleosides **22**, **38**, **95**, **98** and **99** were crystallised and their crystal structures successfully solved for the first time.

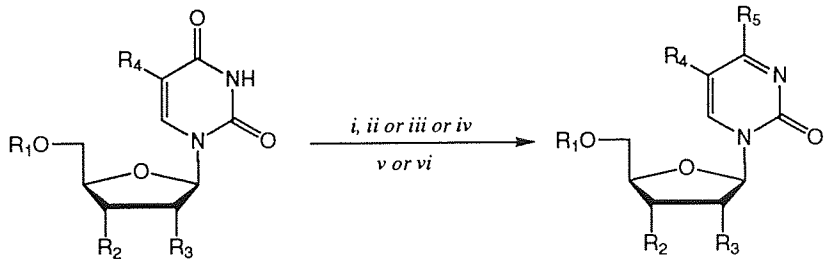


Figure 4.1 Synthesis of modified pyrimidine nucleosides; *i* C_5H_5N; *ii* TMSCl, $ClC_6H_4OP(O)Cl_2$; *iii* $(CF_3CO)_2O$; *iv* $ClC_6H_4OP(O)Cl_2$; *v* C_6F_5OH; *vi* 1,2,4-triazole.

Nucleoside **2** was made in accordance with the literature[119] and was crystallised from ethyl acetate. A better crystal structure refinement was obtained (R = 0.056) compared to the literature report (R = 0.086).[153]

The use of methods A (reagents *i, iii*) and B (reagents *i, ii*) were unsuccessful when applied to 2'-deoxythymidine in the attempted synthesis of nucleosides **92** and **93**. Synthesis of the ribose analogue **22** is described in section (2.4.1). The AZT derivative **99** could only be synthesised using method B whilst nucleoside **100** could not be made using either method. AZT derivative **97** was made successfully by method B, but time did not allow a repeat attempt using method A. Nucleosides **95** and **96** were both synthesised by method A, but again, time did not allow a repeat attempt using method B. The general observation was made that the C4 of both uridine and 2'-deoxyuridine could be activated using

trifluoroacetic anhydride and a C4-substituent, either 1,2,4-triazolyl or pentafluorophenoxy, introduced. When the hydrogen at C5 of the pyrimidine nucleoside was replaced by a methyl substituent, activation at the C4 carbonyl by trifluoroacetic anhydride was apparently hindered.

Starting materials					Products						
R_1	R_2	R_3	R_4	Reagents	No.	R_1	R_2	R_3	R_4	R_5	Yield
H	OH	H	H	*i, ii, v*	**38**	H	OH	H	H	C_6F_5O	78
H	OH	H	H	*i, iii, v*	**38**	H	OH	H	H	C_6F_5O	65
H	OH	H	H	*i, ii, vi*	**90**	H	OH	H	H	Tzl	79
Ac	Ac	H	H	*i, iv, vi*	**22**	Ac	Ac	H	H	Tzl	53
Ac	Ac	H	H	*i, iv, v*	**91**	Ac	Ac	H	H	C_6F_5O	61
H	OH	H	CH_3	*i, ii, v*	**92**	H	OH	H	CH_3	C_6F_5O	-
H	OH	H	CH_3	*i, iii, v*	**92**	H	OH	H	CH_3	C_6F_5O	-
H	OH	H	CH_3	*i, ii, vi*	**93**	H	OH	H	CH_3	Tzl	-
H	OH	H	CH_3	*i, iii, vi*	**93**	H	OH	H	CH_3	Tzl	-
Ac	Ac	H	CH_3	*i, iv, v*	**94**	Ac	Ac	H	CH_3	C_6F_5O	67
Ac	Ac	H	CH_3	*i, iv,vi*	**2**	Ac	Ac	H	CH_3	Tzl	49
H	OH	OH	H	*i, iii, v*	**95**	H	OH	OH	H	C_6F_5O	59
H	OH	OH	H	*i, iii, vi*	**96**	H	OH	OH	H	Tzl	56
Ac	N_3	H	CH_3	*i, iv, v*	**97**	Ac	N_3	H	CH_3	C_6F_5O	62
Ac	N_3	H	CH_3	*i, iv, vi*	**98**	Ac	N_3	H	CH_3	Tzl	88
H	N_3	H	CH_3	*i, ii, v*	**99**	H	N_3	H	CH_3	C_6F_5O	62
H	N_3	H	CH_3	*i, iii, v*	**99**	H	N_3	H	CH_3	C_6F_5O	-
H	N_3	H	CH_3	*i, ii, vi*	**100**	H	N_3	H	CH_3	Tzl	-

Table 4.1 Summary of synthesis details for the preparation of C4-protected pyrimidine nucleosides.

4.3 Biological Results

4.3.1 Anti-HIV data

The C4 pentafluorophenoxy and 1,2,4-triazolyl-substituted pyrimidine nucleosides, along with a number of N4-substituted pyrimidine analogues, were tested for anti-HIV activity using HIV-1 infected C8166 cells (Table 4.3).[154] EC_{50} refers to the concentration of anti-HIV agent required to reduce HIV-I infection by 50% while TC_{50} refers to the concentration required to kill half of healthy, uninfected cells.

Compound	$EC_{50}/\mu M$	$TC_{50}/\mu M$
2	40	>80
22	400	>1000
23	80	>1000
24	8	20
26	>400	>400
38	400	>1000
41	>1000	>1000
42	8	20
90	>1000	>1000
91	100	100
94	>80	80
95	1.6	200
96	100	400
97	2	80
98	1.6	80
99	400	400
AZT	0.016	>1000

Table 4.2 EC_{50} and TC_{50} data for synthesised, C4-substituted pyrimidine nucleosides.

With the exceptions of the N4-substituted nucleoside **41** and 1,2,4-triazolyl nucleoside **90**, all of the compounds displayed anti-HIV activity. However, none of the nucleosides tested were as active as AZT against HIV-1. The azido nucleoside **99** showed only modest anti-HIV activity while the acetyl-protected 1,2,4-triazolyl analogue **98** was more active. Activity was observed for the pentafluorophenoxy-substituted ribose nucleoside **95** which was less cytotoxic than the equipotent azido analogue **98**. While the N4-substituted nucleoside **23** showed only modest activity and was not cytotoxic, its 4,4'-dimethoxytritylated analogue **24** displayed better anti-HIV activity with an EC_{50} of 8 μM and higher cytotoxicity with a TC_{50} of 20 μM. The acetyl-protected 1,2,4-triazolyl nucleoside **22** also showed anti-HIV activity and was not cytotoxic, while its non-acetylated analogue **90** displayed no anti-HIV activity nor cytotoxicity.

A primary hydroxyl is required at the C5' of anti-HIV nucleosides as a substrate for cellular kinase activity if the nucleoside is to act as a RT inhibitor through chain termination (Section 1.4.1). Nucleosides **2**, **91**, **97** and **98** all have C5'-protected hydroxyls but did still display anti-HIV activity. Hydrolysis of the C5'-protecting group, at some stage to unmask the 5'-OH, may account for the observed anti-HIV activity. These compounds would therefore be acting as prodrugs although with greatly reduced activity compared with AZT. Mechanisms other than that of RT-mediated chain termination are of course possible.

Nucleosides **38** and **99** showed anti-HIV activity and both have a free C5' hydroxyl. The reduced activity of nucleoside **99** compared with AZT may reflect a reduced specificity for thymidine kinases which is required for conversion to the triphosphate. The affinity of any of the nucleoside analogue triphosphates for RT may also have been decreased. The protected 1,2,4-triazolyl-substituted AZT analogue **98** may be a prodrug of AZT with the acetyl and 1,2,4-triazolyl functions being hydrolysed within the biological test system.

Any of these compounds could potentially interfere with pyrimidine salvage processes or cause feedback inhibition of pyrimidine biosynthesis. Inhibition of ribonucleotide reductase or nucleoside/nucleotide kinase enzymes is also a possibility. Lack of pronounced anti-viral activity and time constraints did not warrant a thorough investigation of the precise mode of action of these compounds.

4.3.2 Anti-cancer activity

A number of modified pyrimidine nucleosides synthesised in this work were tested as anti-cancer agents by Augustin Dick at Aston University along with AZT for comparison (Table 4.3).[155]

Nucleoside	ID_{50} Values	
	MAC 13	MAC 16
2	inactive	inactive
22	inactive	inactive
38	16.4*	n/a
90	inactive	inactive
91	inactive	inactive
94	36.8	n/a
95	inactive	inactive
96	inactive	inactive
97	55.4	66.1*
98	9.1	24.5
99	inactive	inactive
AZT	3.3	39.3

Table 4.3 Anti-cancer data for modified pyrimidine nucleosides; * = no repeat experiment conducted; $ID_{50}/\mu M$

The modified nucleosides were tested against mouse adenocarcinoma 16 (MAC 16) and mouse adenocarcinoma 13 (MAC 13) cell lines. Four of the modified nucleosides displayed modest anti-cancer activity against MAC 13 and MAC 16 cell lines although none was as active as AZT. Nucleoside **98** showed the greatest activity against the MAC 13 cell line, having an ID_{50} of 9.1 μM.

4.4 Crystallographic studies

The crystal structures of the four C4-modified pyrimidine nucleosides **22, 38, 95 98** and **99** have been successfully solved and their crystallographic parameters summarised in Table 4.9, 4.11, 4.13, 4.15 and 4.17 respectively. The known nucleoside **2** has been refined to give a superior *R* factor compared with the literature[153] (Table 4.7) and one of the known AZT structures has been included for comparison.[156] The crystal structure of **98** shows the glycosidic bond close to an eclipsed conformation with the sugar ring out of plane with C3' *endo* and C4' *exo*. The triazole ring is conjugated with the pyrimidine ring and the azido group is almost perpendicular to the C2'-C3' bond and not *trans* to it as in AZT. In the crystal structure of nucleoside **99**, the bridging oxygen at O4 is less conjugated with the pentafluorophenoxy ring than with the pyrimidine, which is oriented *anti* to the sugar with its C2' *endo*, C3' *exo* pucker and *gauche-gauche* 5'-OH group. Similarly, the crystal structure of compound **38** exhibits a C2' *endo*, C3' *exo* sugar pucker with the O4 less conjugated with the pentafluorophenoxy ring than with the pyrimidine.

The N2" nitrogen of the 1,2,4-triazole ring is *anti* with respect to the N2 of the pyrimidine ring in nucleosides **22** and **98**. This may be due to 1,3-repulsion between the non-bonding lone pair of electrons of the N2 and the N2". The plane of the pentafluorophenoxy ring of nucleosides **38, 95, 99** is consistently perpendicular with the pyrimidine ring. It is thought that one of the non-bonding lone pair of electrons of the O4 oxygen is conjugated more with the pyrimidine ring than with the pentafluorophenoxy ring thus reducing rotation around the C4-O4 bond. It is also possible that the lone pair of electrons of N3 can interact with the electron-deficient pentafluorophenoxy ring above and that the pentafluorophenoxy

rings of adjacent molecules in the crystal lattice stacked one on top of the other, presumably due to favourable σ-π interactions. Additionally, rotation around the bond connecting the O4 to the pentafluorophenoxy ring (O4-C1") is possibly restricted due to N3-F1 and N3-F6 1,5-repulsions.

Nucleoside	*syn or anti*	χ_{CN}	Sugar Conformation
AZT	*anti*	52.9	C2'-*endo*
1	*anti*	50.67	C2'-*endo*
22	*anti*	12.18	C2'-*endo*
38	*anti*	52.40	C2'-*endo*
95	*anti*	54.49	C2'-*endo*
98	*anti*	15.79	C2'-*endo*
99	*anti*	21.37	C3'-*endo*

Table 4.4 Heterocycle orientation, glycosidic torsion angle and sugar conformation.

4.5 Conclusion

Members of a series of nucleoside analogues, synthesised in good yields by direct methods, were shown to possess anti-HIV activity. None of these compounds were as active as AZT and none were without cytotoxicity. Some members of the series also displayed anti-cancer activity.

Time constraints and lack of profound biological activity did not permit a more extensive study of structure activity relationships although the structural information uncovered from the crystal structure determinations could provide a starting point for the design of more potent anti-viral compounds.

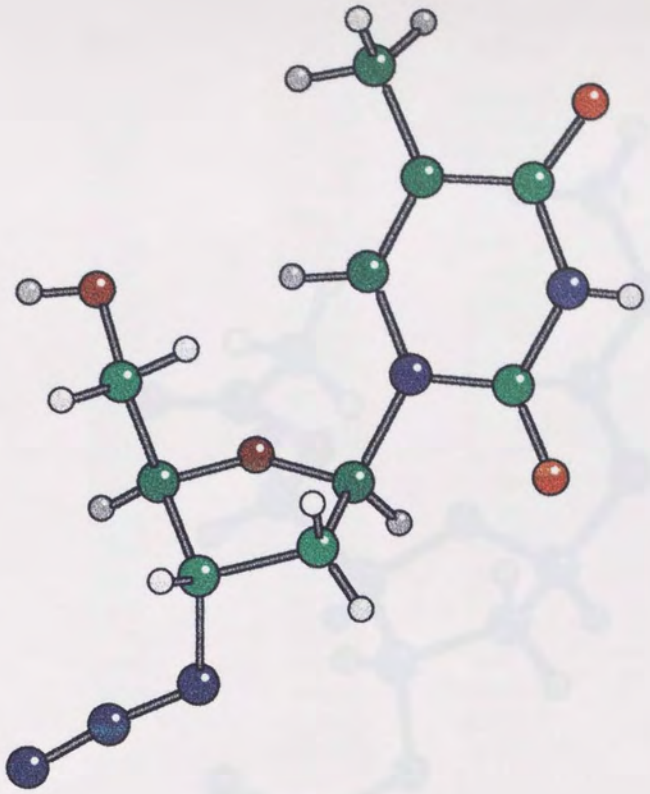


Figure 4.2 Ball-and-stick representation of nucleoside AZT.

Compound	**AZT**	**Cell Parameters /Å** *a*	5.628(4)
Solvent	$EtOH/Me_2O$	*b*	12.013(7)
Formula	$C_{10}H_{13}N_5O_4$	*c*	17.507(10)
System	Monoclinic	**Volume (*V* /Å^3)**	1608.3
Space Group	$P2_1$	**Z**	4
R	0.028	β (°)	95.96(5)

Table 4.5 Crystallographic parameters for nucleoside AZT.[153]

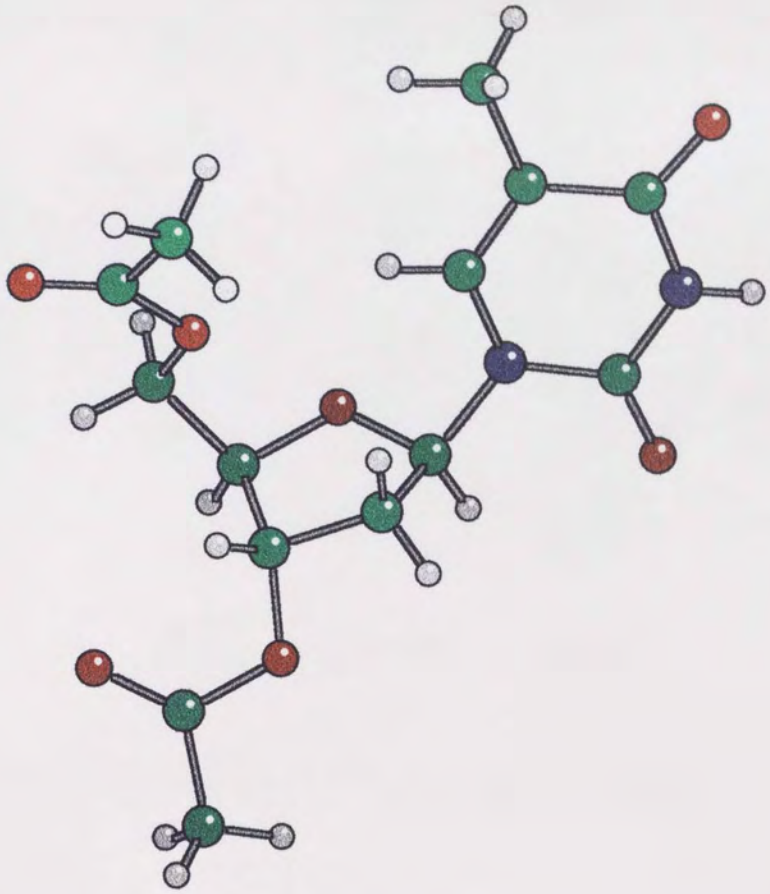


Figure 4.3 Ball-and-stick representation of nucleoside **2**.

Compound	**2**	**Cell Parameters /Å** *a*	7.074(2)
Solvent	EtOAc	*b*	8.843(2)
Formula	$C_{14}H_{17}N_2O_7$	*c*	9.819(2)
System	Monoclinic	**Volume (*V*/Å^3)**	1608.3
Space Group	$P2_1$	**Z**	4
R	0.0564	β (°)	100.68(1)

Table 4.6 Crystallographic parameters for nucleoside **2**.

Atom	x	y	z	U(eq)
N(1)	3357(2)	4641(1)	4834(2)	46(1)
C(2)	4410(2)	4813(1)	4008(2)	47(1)
N(3)	3745(2)	5142(1)	2782(2)	49(1)
C(4)	2219(3)	5308(1)	2330(2)	55(1)
C(5)	1189(3)	5122(2)	3266(3)	61(1)
C(6)	1833(3)	4801(1)	4462(3)	55(1)
O(7)	5763(2)	4698(1)	4308(2)	65(1)
O(8)	1809(2)	5606(1)	1210(2)	76(1)
C(9)	-491(4)	5298(3)	2857(4)	100(1)
C(1')	3941(2)	4343(1)	6218(2)	44(1)
C(2')	3822(3)	4858(1)	7378(2)	49(1)
C(3')	3538(3)	4384(1)	8550(2)	50(1)
C(4')	2725(3)	3738(1)	7808(2)	48(1)
O(4')	2972(2)	3762(1)	6396(1)	46(1)
O(3')	5012(2)	4205(1)	9364(2)	57(1)
C(6')	4985(3)	3849(2)	10568(2)	63(1)
O(6')	3830(3)	3673(2)	10909(2)	83(1)
C(7')	6572(4)	3742(2)	11362(3)	84(1)
C(5')	1036(3)	3721(1)	7827(3)	63(1)
O(5')	378(2)	4395(1)	7320(2)	70(1)
C(8')	-942(4)	4590(2)	7658(5)	92(1)
O(8')	-1581(7)	4227(3)	8327(9)	227(4)
C(9')	-1488(4)	5292(2)	7053(6)	98(1)

Table 4.7 Atomic co-ordinates (x 10^4) and equivalent isotropic displacement parameters ($Å^2$ x 10^3) for **2**.

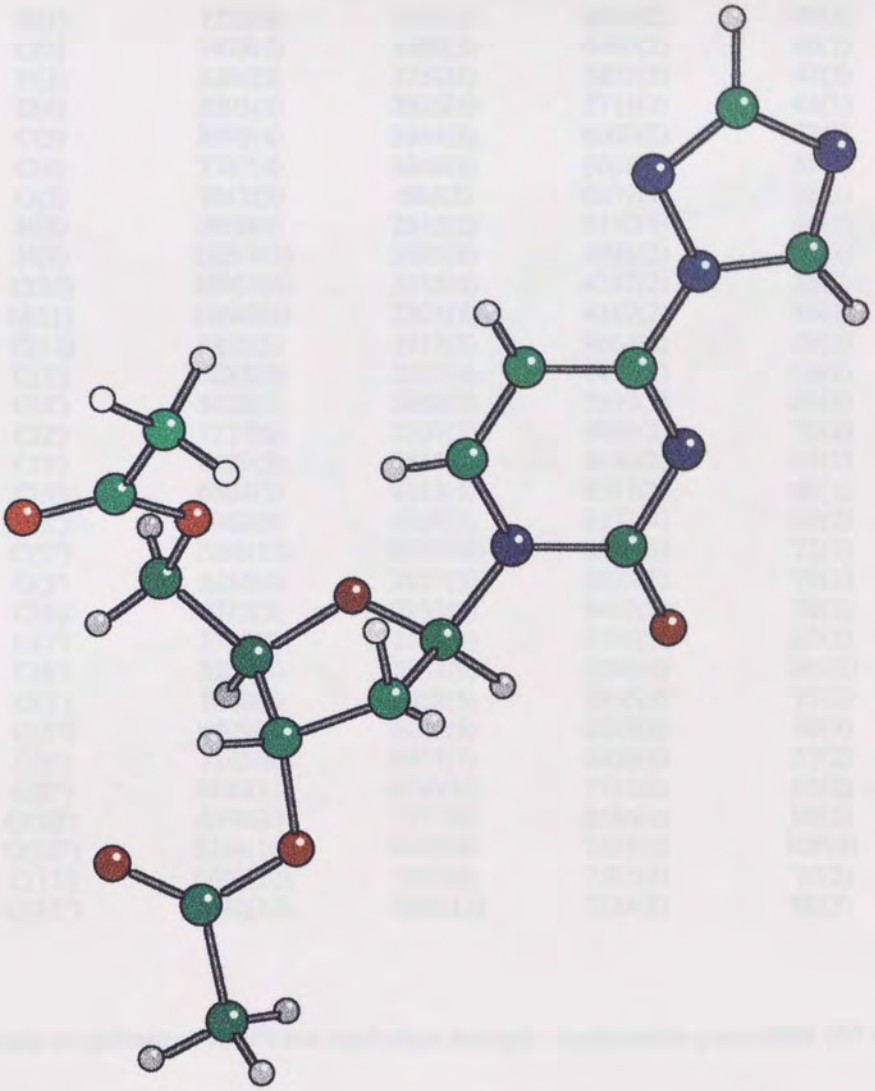


Figure 4.4 Ball-and-stick representation of nucleoside **22**.

Compound	**22**	**Cell Parameters/Å** *a*	8.417(3)
Solvent	EtOAc	*b*	10.824(6)
Formula	$C_{15}H_{16}N_5O_6$	*c*	18.885(5)
System	Monoclinic	**Volume** (V/Å^3)	1700.7
Space Group	$P2_1$	**Z**	4
R	0.0557	β(°)	98.70(2)

Table 4.8 Crystallographic parameters for nucleoside **22.**

Atom	x	y	z	U(eq)
N(1)	7222(3)	2766(3)	6824(2)	49(1)
C(2)	7478(3)	1666(3)	6460(2)	46(1)
N(3)	8266(3)	1752(2)	5877(2)	47(1)
C(4)	8795(3)	2825(3)	5711(2)	44(1)
C(5)	8599(4)	3944(3)	6063(2)	52(1)
C(6)	7767(4)	3868(3)	6618(2)	55(1)
O(7)	7012(3)	684(2)	6677(1)	58(1)
N(8)	9600(3)	2845(2)	5110(1)	46(1)
N(9)	10291(3)	3895(3)	4901(2)	54(1)
C(10)	10901(4)	3515(4)	4342(2)	56(1)
N(11)	10640(4)	2305(3)	4167(2)	66(1)
C(12)	9823(5)	1917(3)	4664(2)	59(1)
C(1')	6283(4)	2665(4)	7423(2)	58(1)
O(4')	5820(3)	3862(3)	7599(1)	65(1)
C(2')	7217(6)	2107(5)	8099(2)	70(1)
C(3')	6689(5)	2863(4)	8696(2)	61(1)
C(4')	6304(5)	4113(4)	8341(2)	69(1)
C(5')	7962(9)	4869(7)	8371(4)	69(2)
C(5")	7293(13)	5090(10)	8589(6)	75(3)
O(3')	5282(4)	2273(3)	8877(2)	78(1)
C(6')	4715(5)	2732(5)	9453(2)	78(1)
O(7')	5316(4)	3570(4)	9796(2)	83(1)
C(8')	3269(9)	2041(9)	9596(4)	145(3)
O(5')	7848(7)	5852(5)	7856(3)	75(2)
O(5")	6629(9)	6186(8)	8243(6)	98(4)
C(9')	7165(8)	6927(7)	8005(4)	57(2)
C(9")	7436(11)	6736(10)	7791(6)	62(2)
O(10')	6976(8)	7175(8)	8586(4)	96(2)
O(10")	8744(10)	6420(9)	7693(6)	109(4)
C(11')	6697(10)	7647(9)	7305(5)	70(2)
C(11")	6582(14)	7885(12)	7534(8)	80(3)

Table 4.9 Atomic co-ordinates (x 10^4) and equivalent isotropic displacement parameters ($Å^2$ x 10^3) for **22**.

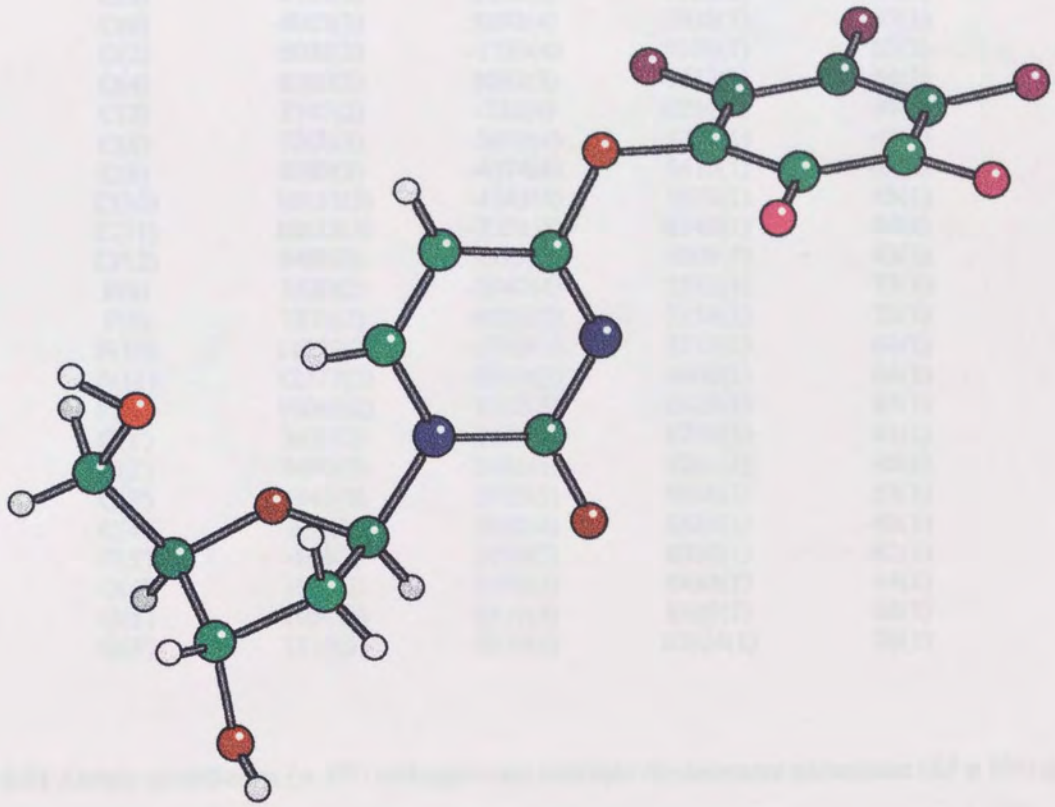


Figure 4.5 Ball-and-stick representation of nucleoside **38**.

Compound	**38**	**Cell Parameters/Å** *a*	7.739(2)
Solvent	Toluene/MeOH	*b*	5.502(1)
Formula	$C_{15}H_{11}F_5N_2O_5$	*c*	18.546(2)
System	Monoclinic	**Volume (*V*/Å^3)**	780.9
Space Group	$P2_1$	***Z***	2
R	0.0340	β(°)	98.54(2)

Table 4.10 Crystallographic parameters for nucleoside **38**.

Atom	x	y	z	U(eq)
N(1)	4392(2)	1476(3)	8147(1)	40(1)
C(2)	5640(2)	-320(4)	8095(1)	42(1)
N(3)	6367(2)	-436(3)	7463(1)	42(1)
C(4)	5930(2)	1196(4)	6973(1)	40(1)
C(5)	4782(3)	3153(4)	7016(1)	46(1)
C(6)	4005(3)	3192(4)	7618(1)	43(1)
O(2)	6036(2)	-1753(4)	8598(1)	63(1)
O(4)	6590(2)	1092(3)	6332(1)	54(1)
C(7)	7747(2)	-752(4)	6231(1)	44(1)
C(8)	7208(3)	-2693(4)	5790(1)	47(1)
C(9)	8389(3)	-4374(4)	5610(1)	48(1)
C(10)	10133(3)	-4143(4)	5893(1)	45(1)
C(11)	10683(3)	-2221(4)	6345(1)	44(1)
C(12)	9496(3)	-535(4)	6509(1)	43(1)
F(8)	5520(2)	-2947(4)	5512(1)	73(1)
F(9)	7876(2)	-6215(3)	5158(1)	70(1)
F(10)	11280(2)	-5749(3)	5711(1)	64(1)
F(11)	12377(2)	-2003(3)	6606(1)	64(1)
F(12)	10040(2)	1352(3)	6929(1)	63(1)
C(1')	3410(2)	1460(4)	8770(1)	41(1)
C(2')	3693(3)	3686(4)	9261(1)	48(1)
C(3')	1941(3)	3915(5)	9536(1)	53(1)
C(4')	635(3)	3082(4)	8884(1)	49(1)
C(5')	-198(3)	5066(5)	8375(1)	62(1)
O(4')	1607(2)	1441(3)	8483(1)	44(1)
O(5')	1051(3)	6516(4)	8101(1)	66(1)
O(3')	1813(3)	2233(5)	10104(1)	76(1)

Table 4.11 Atomic co-ordinates (x 10^4) and equivalent isotropic displacement parameters ($Å^2$ x 10^3) for **38**.

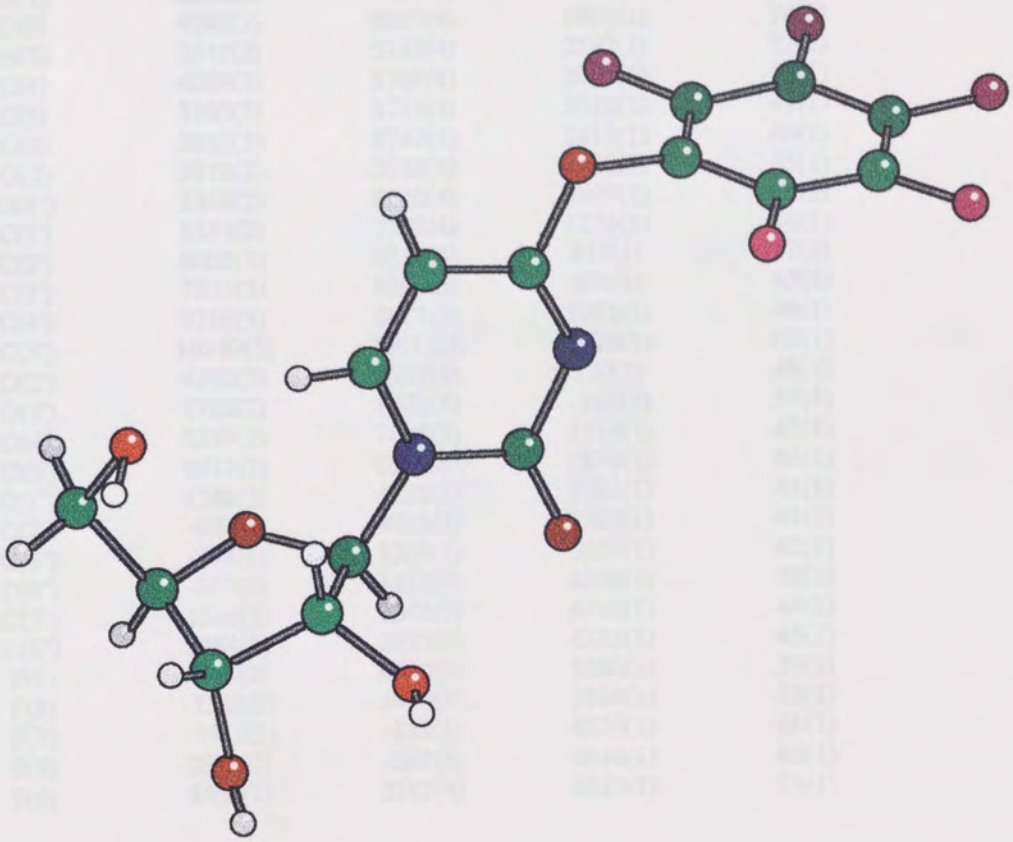


Figure 4.6 Ball-and-stick representation of nucleoside **95**.

Compound	**95**	**Cell Parameters/Å** *a*	7.564(1)
Solvent	MeOH	*b*	5.561(1)
Formula	$C_{15}H_{11}F_5N_2O_5$	*c*	19.031(2)
System	Monoclinic	**Volume (*V*/Å^3)**	791.6
Space Group	$P2_1$	***Z***	2
R	0.0286	*β*(°)	98.54(6)

Table 4.12 Crystallographic parameters for nucleoside **95**.

Atom	x	y	z	U(eq)
N(1)	5509(2)	7043(4)	1904(1)	36(1)
C(2)	4248(3)	5253(4)	1963(1)	38(1)
N(3)	3545(2)	5145(4)	2592(1)	37(1)
C(4)	4009(2)	6769(4)	3063(1)	36(1)
C(5)	5160(3)	8715(4)	3010(1)	41(1)
C(6)	5920(3)	8747(4)	2412(1)	40(1)
O(2)	3810(2)	3833(4)	1480(1)	55(1)
O(4'')	3368(2)	6660(4)	3699(1)	49(1)
C(1')	6433(2)	7072(4)	1274(1)	36(1)
C(2')	6065(3)	9315(4)	817(1)	37(1)
C(3')	7803(3)	9562(5)	494(1)	45(1)
C(4')	9216(3)	8671(5)	1091(1)	46(1)
C(5')	10149(3)	10611(6)	1576(1)	62(1)
O(2')	4502(2)	8927(4)	320(1)	49(1)
O(3')	7788(2)	8028(4)	-101(1)	57(1)
O(4')	8299(2)	7044(3)	1518(1)	42(1)
O(5')	8917(3)	12073(4)	1870(1)	64(1)
C(1'')	2188(3)	4828(5)	3791(1)	41(1)
C(2'')	409(3)	4993(5)	3507(1)	41(1)
C(3'')	-797(3)	3307(5)	3659(1)	42(1)
C(4'')	-237(3)	1433(5)	4108(1)	43(1)
C(5'')	1536(3)	1242(5)	4398(1)	44(1)
C(6'')	2744(3)	2927(5)	4235(1)	45(1)
F(1)	-145(2)	6840(3)	3080(1)	59(1)
F(2)	-2523(2)	3490(3)	3384(1)	63(1)
F(3)	-1411(2)	-181(3)	4276(1)	61(1)
F(4)	2063(2)	-569(3)	4844(1)	66(1)
F(5)	4461(2)	2715(4)	4517(1)	71(1)

Table 4.13 Atomic co-ordinates (x 10^4) and equivalent isotropic displacement parameters ($Å^2$ x 10^3) for **95**.

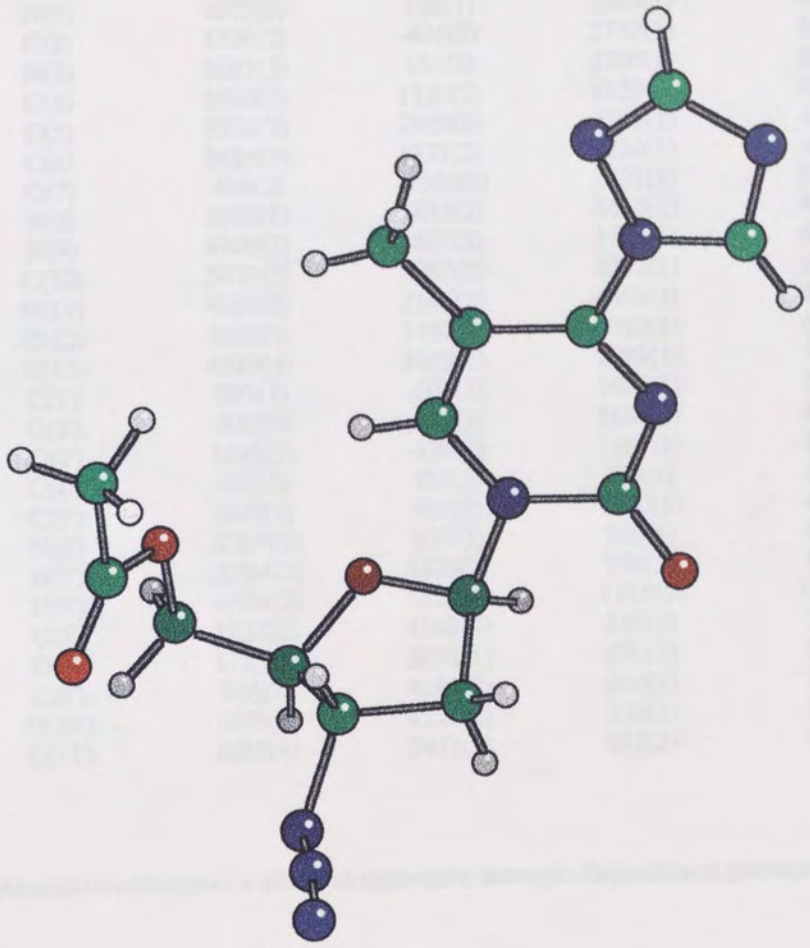


Figure 4.7 Ball-and-stick representation of nucleoside **98**.

Compound	**98**	**Cell Parameters/Å** *a*	8.513(3)
Solvent	EtOAc/MeOH	*b*	8.720(3)
Formula	$C_{14}H_{16}N_8O_4$	*c*	22.855(6)
System	Orthorhombic	**Volume** (V/Å^3)	1696.4
Space Group	$P2_12_12_1$	Z	4
R	0.0263	β(°)	90

Table 4.14 Crystallographic parameters for nucleoside **98**.

Atom	x	y	z	U(eq)
N(1)	1679(2)	148(1)	2189(1)	44(1)
C(2)	1356(2)	-476(2)	2737(1)	53(1)
N(3)	2097(2)	151(2)	3205(1)	50(1)
C(4)	3046(2)	1320(2)	3129(1)	42(1)
C(5)	3374(2)	2029(2)	2591(1)	45(1)
C(6)	2633(2)	1371(2)	2130(1)	44(1)
O(7)	446(2)	-1565(2)	2771(1)	80(1)
N(8)	3698(2)	1835(2)	3665(1)	47(1)
N(9)	4909(2)	2867(2)	3708(1)	60(1)
C(10)	5073(3)	2987(2)	4272(1)	64(1)
N(11)	4092(2)	2146(2)	4604(1)	72(1)
C(12)	3248(3)	1437(2)	4210(1)	63(1)
C(13)	4389(4)	3419(3)	2488(1)	74(1)
C(1')	903(3)	-606(2)	1685(1)	55(1)
C(2')	-822(3)	-187(3)	1633(1)	64(1)
O(4')	1648(2)	-126(1)	1167(1)	54(1)
C(4')	491(2)	421(2)	758(1)	47(1)
C(3')	-848(2)	964(2)	1140(1)	46(1)
N(6')	-2299(2)	933(2)	788(1)	66(1)
N(7')	-3384(2)	1676(2)	986(1)	68(1)
N(8')	-4474(3)	2322(4)	1115(1)	113(1)
C(5')	1222(2)	1568(2)	359(1)	53(1)
O(5')	1753(1)	2898(1)	676(1)	56(1)
C(9')	908(2)	4193(2)	616(1)	58(1)
O(10')	-263(2)	4269(2)	333(1)	72(1)
C(11')	1621(4)	5471(3)	952(2)	96(1)

Table 4.15 Atomic co-ordinates (x 10^4) and equivalent isotropic displacement parameters ($Å^2$ x 10^3) for **98**.

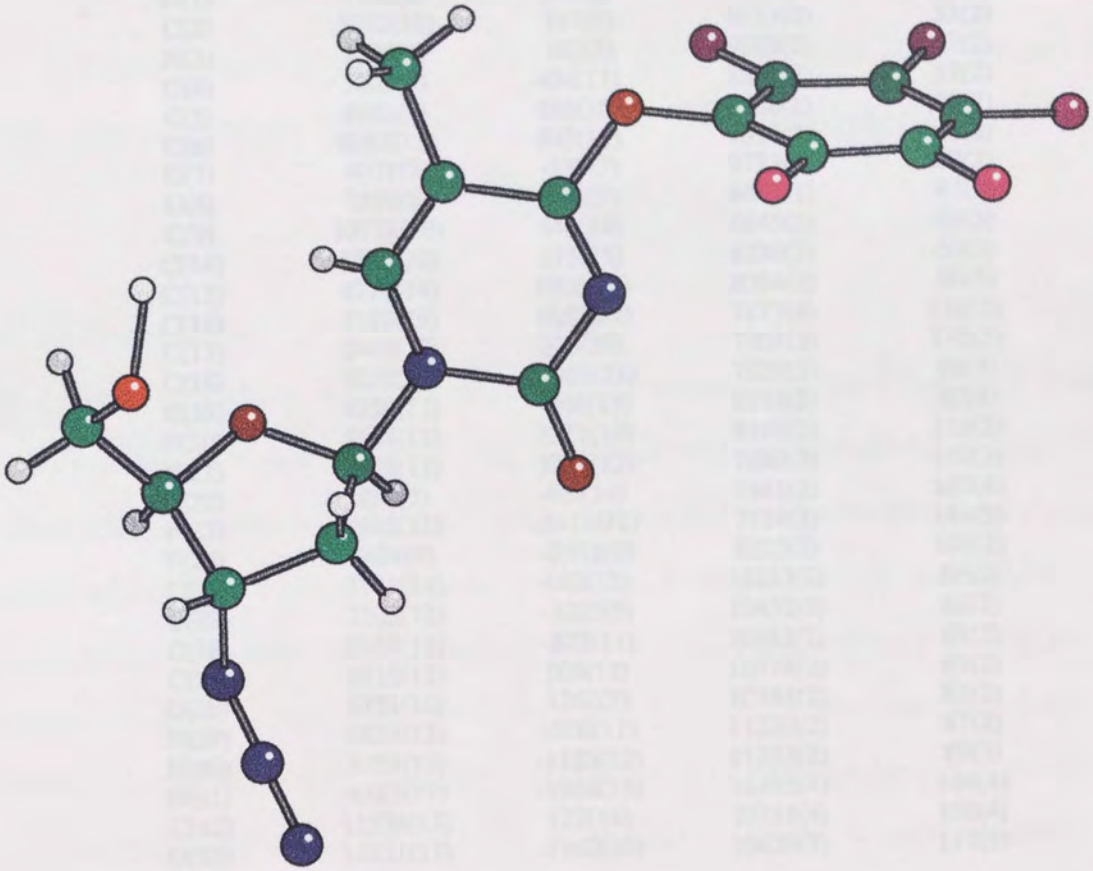


Figure 4.8 Ball-and-stick representation of nucleoside **99**.

Compound	**99**	**Cell Parameters /Å** *a*	7.074(2)
Solvent	EtOAc/MeOH	*b*	8.197(4)
Formula	$C_{16}H_{12}F_5N_5O_4$	*c*	30.588(11)
System	Orthorhombic	**Volume (*V*/Å^3)**	1773.7
Space group	$P2_12_12_1$	**Z**	4
R	0.0564	β(°)	90

Table 4.16 Crystallographic parameters for nucleoside **99**.

Atom	x	y	z	U(eq)
N(1)	7238(8)	379(8)	9757(2)	55(2)
C(2)	5532(10)	117(9)	9533(2)	53(2)
N(3)	5596(8)	160(8)	9089(2)	57(2)
C(4)	7203(9)	404(11)	8891(2)	57(2)
C(5)	8982(9)	583(10)	9108(2)	60(2)
C(6)	8882(11)	545(12)	9537(2)	60(2)
O(7)	4071(7)	-137(7)	9733(2)	69(2)
O(8)	7256(6)	498(9)	8447(1)	82(2)
C(9)	10778(10)	848(14)	8845(3)	89(3)
C(14)	5586(10)	513(15)	8230(2)	65(2)
C(15)	4772(14)	1906(17)	8084(3)	80(3)
C(16)	3182(19)	1850(24)	7827(4)	116(5)
C(17)	2447(12)	326(30)	7701(3)	116(5)
C(18)	3220(15)	-1008(21)	7829(3)	98(4)
C(19)	4750(13)	-946(15)	8102(3)	80(3)
F(20)	5564(11)	3321(10)	8190(2)	116(2)
F(21)	2455(11)	3294(12)	7680(3)	167(3)
F(22)	896(7)	409(14)	7441(2)	167(4)
F(23)	2415(11)	-2419(12)	7714(2)	144(3)
F(24)	5481(9)	-2355(9)	8225(2)	109(2)
C(25)	7141(14)	442(12)	10233(2)	66(2)
C(29)	7205(12)	-1225(9)	10452(3)	66(2)
C(28)	8168(11)	-875(11)	10881(2)	63(2)
C(27)	9513(12)	509(12)	10774(2)	67(2)
O(26)	8751(10)	1262(7)	10384(2)	87(2)
N(39)	6853(12)	-206(11)	11220(2)	87(2)
N(40)	5755(13)	-1183(12)	11353(2)	89(3)
N(41)	4587(17)	-1984(15)	11495(4)	144(4)
C(42)	11554(15)	122(14)	10711(4)	108(4)
O(33)	11811(11)	-1160(10)	10429(3)	117(3)

Table 4.17 Atomic co-ordinates (x 10^4) and equivalent isotropic displacement parameters ($Å^2$ x 10^3) for **99**.

Chapter 5

A New Cleavage/Deprotection Strategy for Automated Oligonucleotide Synthesis Using a Novel Silyl-Linked Solid Support

5.1 Introduction

The general applicability of phosphodiester-linked antisense oligonucleotides (O-oligos) as therapeutic agents is problematic. O-Oligos are susceptible to nuclease hydrolysis and have a short half-life of only a few hours in calf serum. The cellular membrane acts as a physical barrier when challenged by negatively charged O-oligos; the estimated intracellular concentration of O-oligos being between 7.5-10% of the extracellular concentration.[157] A number of chemical modifications to the O-oligo backbone has been introduced in an attempt to improve the half-life and cellular uptake of oligonucleotides. Phosphorothioate oligonucleotides (S-oligos)[158,159] and methylphosphonate oligonucleotides (MP-oligos)[160,161] have been extensively studied due to their improved resistance to nuclease hydrolysis. Substitution of one of the diastereotopic oxygens of the O-oligo internucleotide linkage by sulphur (S-oligos) or methyl (MP-oligos) leads to the formation of a chiral centre at phosphorus (Figure 5.1). The pH of extracellular fluids confers a negatively charged backbone on O- and S-oligos while MP-oligos remain neutral. In the case of S-oligos, the negative charge has been shown the reside on sulphur.[162,163] MP-, S- and O-oligos have been shown to enter cells and have proved to be useful tools as genetic probes in helping to elucidate the general mechanism of cellular uptake and intracellular distribution of oligonucleotides.

5.1.1 Cellular uptake of phosphodiester and phosphorothioate oligonucleotides

Uptake of oligonucleotides into cells *in vitro* is affected by a number of factors (Table 5.1).[164] Care is always needed when comparing results from different sources since

different cell lines may express different types and numbers of cell surface proteins. Oligonucleotide labels, such as covalently linked hydrophobic fluorescent probes used to visualise the intracellular distribution of oligonucleotides, may also affect cellular uptake.

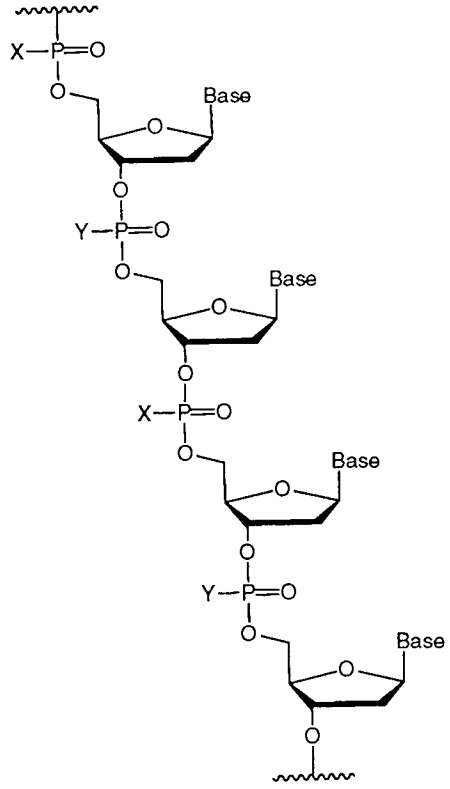


Figure 5.1 Backbone modifications to oligonucleotides: O-Oligos ($X=O^-$, $Y=O^-$); S-oligos ($X=S^-$, $Y=S^-$); MP-Oligos ($X=CH_3$, $Y=CH_3$) and S-O ($X=O^-$, $Y=S^-$), MP-O-oligos ($X=O^-$, $Y=CH_3$) chimerae.

Factor	Comment
Type of cell	Number and type of protein receptor
Components of media	5'-Nucleotide monophosphates (e.g. AMP) compete for cellular uptake; nucleases cleave O-oligos
Type of backbone	Passive or active transport
Length of oligomer	Longer the oligomer, slower the uptake
Auxiliary groups	Uptake and stability may be affected by the attached auxiliary

Table 5.1 Possible factors influencing the cellular uptake of oligonucleotides.

The polyanionic character of O-oligos suggest that passive diffusion through a cellular lipid membrane would be an improbable mechanism for their cellular uptake. It was not until 1978 that Zamecnik and co-workers showed that O-oligos could enter cells and consequently exert a biological effect.[165,166] Two groups, Loke *et al* (1989)[167] and Yakubov *et al* (1989),[168] suggested that one possible mechanism of cellular uptake of O-oligos was *via* a specific cell surface receptor which might mediate endocytosis (receptor-mediated endocytosis, RME). The cellular uptake process of labelled O-oligos was found to be temperature and energy dependant, saturable and could be inhibited by unlabelled oligonucleotides.

Yakubov and co-workers used an O-oligo conjugated at the 5'-terminus with the reactive group 4-(*N*-2-chloroethyl-*N*-methyl)aminobenzylamine to alkylate the protein thought to be involved in receptor-mediated endocytosis. Two labelled proteins of 79 kDa and 90 kDa which could bind O-oligos were isolated using the mouse fibroblast L929 cell line having a density of 120,000 receptors per cell. S-oligos have been shown to compete for cellular uptake with O-oligos, suggesting uptake occurs *via* a common mechanism of receptor-mediated endocytosis.

Yakubov suggested that fluid-phase endocytosis (FPE, pinocytosis) also has a role in cellular uptake at high oligonucleotide concentrations while adsorptive endocytosis (AE) is responsible for cellular uptake of oligonucleotides at low concentrations (Figure 5.2). Stein *et al* (1993)[169] showed that intracellular accumulation of O- and S-oligos by HL60 cells was concentration and time-dependant, consistent with a pinocytotic mechanism of cellular uptake. S-oligos and, to a lesser extent, O-oligos were shown to inhibit the rate of internalisation of fluorescent albumin; a marker for pinocytosis.

5.1.2 Cellular uptake of methylphosphonate oligonucleotides

MP-oligos, having a neutral backbone and therefore greater hydrophobicity, were thought to be capable of passive diffusion across lipid membranes. Shoji *et al* (1991)[170] showed

that MP-oligo uptake by CH^RC5 cells was temperature-dependent, suggesting that an active cellular process was involved. Cellular uptake of MP-oligos was not significantly affected when in competition with excess ATP. Only modest competition was observed between MP-oligos and the corresponding O-oligo, in large excess, suggesting that O-oligos do not effectively compete for the MP-oligo cellular uptake mechanism. The authors felt that an active endocytotic process was involved, possibly adsorptive endocytosis. Zhao *et al* (1993)[171] used spleen cell cultures from six-week-old DBA/2 mice to determine an order for cellular binding and uptake. S-Oligos were shown to have the highest cell binding and uptake, followed by S-O-, O- and MP-O-oligomer chimerae.

5.1.3 Intracellular fate of oligonucleotides

Whichever endocytotic process is involved in cellular uptake of oligonucleotides, to exert their biological effect, efflux form cellular endosomes is necessary. The lipid composition of the endosome membrane is different to that of the cell membrane. It is considered that passive diffusion of oligonucleotides across endosome membranes is possible but this process would be very slow.[172] It is also possible that leakage of oligonucleotides could occur during membrane disruption when other vesicles, such as lysosomes, fuse to the endosomes.[173] Transient membrane destabilisation by membrane proteins allowing oligonucleotide efflux is also thought to play a role.[174,175] Once efflux from the endosome has occurred, the distribution of oligonucleotides within the cell remains unclear. Researchers have shown that oligonucleotides accumulate both in the cytoplasm[176-179] and nucleus of the cell. Micro-injection of oligonucleotides into cells bypasses endocytosis showing localisation of oligonucleotides within the cell nucleus[178,179] (Figure 5.2).

MP- and S-Oligos are stable to nuclease hydrolysis although not totally chemically inert. The stability of O-oligos depends on the cell culture used. For example, a 5'-phosphorylated 15mer was shown to be stable for 24 to 48 hours in a culture of HL-60[180] leukaemia cells whereas a 20mer suffered extensive hydrolysis in HeLa cells.[181] O-Oligos have been shown to be relatively stable in T-lymphocyte cells.[182] It has also been shown

that the rate of enzymatic hydrolysis depends on the cellular compartment. There is rapid degradation of O-oligos in liposomes while their hydrolysis in the cytoplasm is slower.[183]

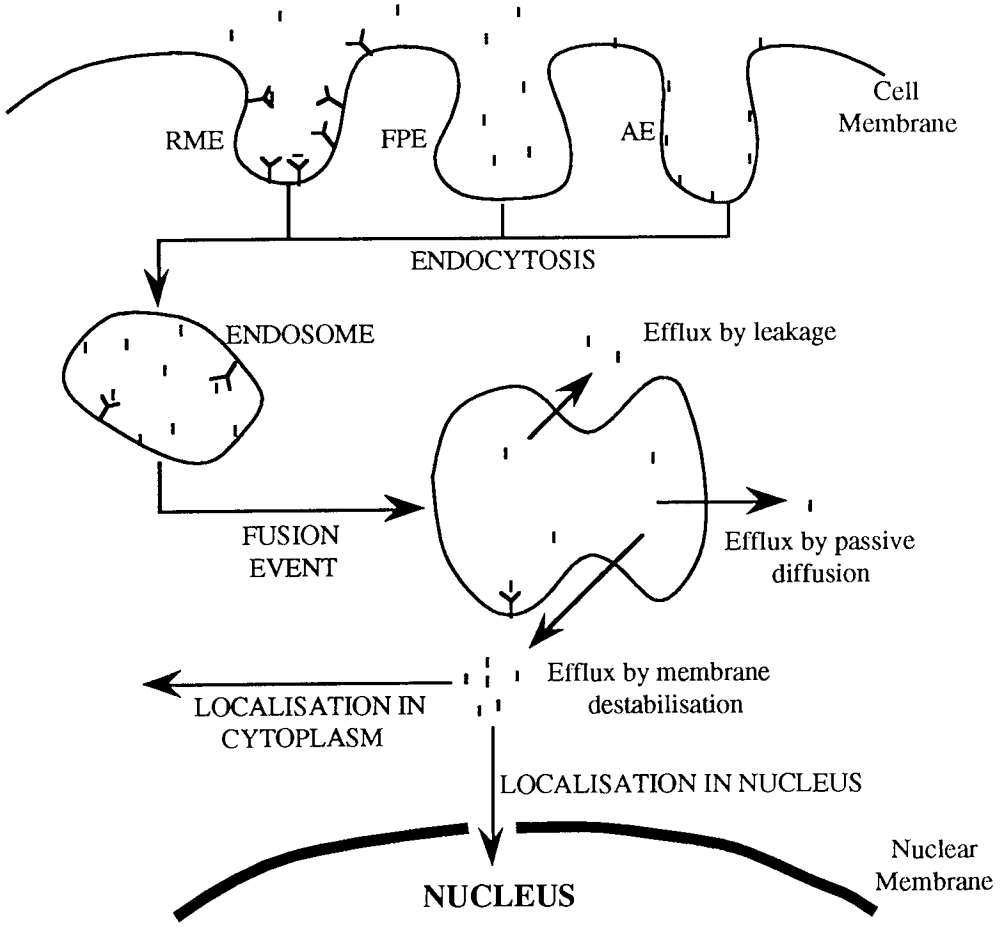


Figure 5.2 Uptake and intracellular fate of oligonucleotides; RME = receptor-mediated endocytosis, FPE = fluid-phase endocytosis and AE = adsorptive endocytosis.

5.2 Limitations of Phosphorothioate and Methylphosphonate Oligos

Although S- and MP-oligos have been shown to have greater nuclease resistance and are capable of entering cells, a number of inherent features of these chemically modified oligonucleotides limit their usefulness.

5.2.1 Non-specific effects

Gao *et al* (1992)[184] used a number of chimeric S-O-oligos having the general formula $d(Cs_xp_{(27-x)})C$ to study the position and number of phosphorothioate linkages on the inhibition of human DNA polymerase and RNase-H activity. S-O-Oligos were shown to inhibit human α-, β- and γ-DNA polymerase and RNase-H *in vitro* in a sequence-independent manner at nanomloar concentrations. This phenomenon is not observed with O-oligos. It was found that this inhibitory effect of S-O-oligos was independent of position and required a minimum of 15 phosphorothioate linkages with a plateau of 28. This suggests the formation of a strong S-O-oligo-enzyme complex due to the electronic nature of the phosphorothioate linkage. Below 15 phosphorothioate centres, the chimeric S-O-oligos had a moderately enhanced ability to induce RNase-H activity. MP-Oligos have no ability to induce RNase-H cleavage of target mRNA.[185] S-Oligos have also been shown to have other non-specific effects, for example, inhibition of HIV transcriptase *in vitro* in a sequence-independent manner.[186]

5.2.2 Chirality at phosphorus

S- and MP-oligos are chiral at phosphorus while O-oligos are achiral. Non-stereospecific synthesis of S- and MP-oligos of length n will yield 2^{n-1} stereoisomers, only a proportion of which will have the correct stereochemistry for optimum hybridisation with target nucleic acid sequences (Figure 5.3).

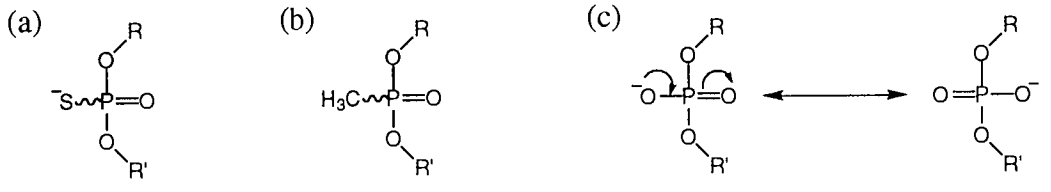


Figure 5.3 Phosphorothioate (a) and methylphosphonate (b) centres are chiral while phosphodiester centres (c) are achiral.

The absolute configuration for phosphorothioate internucleotide linkages are designated Rp or Sp. It has been shown that P-S$^-$ bond in the Rp configuration is oriented into the major groove of duplex DNA.[187,188] LaPlanche *et al* (1986)[189] separated the two distereoisomers of the sequence d(GGsAATTCC) which has one phosphorothioate linkage between A and G residues. Thermal analysis to determine the T_m values of the unmodified phosphodiester parent sequence and the two isomerically pure modified sequences, gave the following order of duplex stability: parent = SpSp > RpRp, with a drop of 2.4°C over the parent sequence. Bower *et al* (1987)[190] stereospecifically replaced, at various positions, phosphodiester linkages for methylphosphonate linkage, in the self-complementary sequence d(GGAATTCC). The RpRp isomeric duplexes have similar T_m values to the parent sequence, with the T_m of the SpSp isomer being lower by between 7 and 11°C.

5.3 New Strategies for Optimising the Properties of Oligonucleotides

Susceptibility to nuclease enzymes presents a major obstacle to using of O-oligos as therapeutic agents. However, within the cell, O-oligos clearly have advantages over S- and MP-oligos. O-oligos are achiral, hybridise with messenger RNA with a higher melting temperature than S- and MP-oligos and are the most efficient at inducing RNase-H activity. Protection of the phosphate centres by groups labile to intracellular enzymes could potentially address the problem of susceptibility of O-oligos to degredation by extracellular nuclease enzymes. Once internalised within a cell, unmasking of the phosphate centres by intracellular enzymes, for example carboxyesterases, would yield O-oligos which could exert antisense or antigene effects. Such phosphate-protected prooligonucleotides might resist nuclease degradation as nuclease activity seems to be sensitive to modifications directly adjacent to the phosphorus centre (Figure 5.4). Uncharged prooligonucleotides might be internalised by cells in a similar way to MP-oligos since uncharged MP-oligos are effectively taken up into cells by a mechanism thought to be distinct from that of negatively charged O-oligos. Such phosphate protected nucleosides could potentially utilise this MP-oligo uptake mechanism. Once internalised, the prooligonucleotide could be deprotected by intracellular enzymes to yield achiral O-oligos.

Acyloxymethyl protection of negatively charged phosphorothioate centres has been used to make short, chiral prooligonucleotides.[191] Once internalised, carboxyesterase-catalysed hydrolysis of the prooligonucleotide yields the unprotected S-oligo. Automated solid-phase synthesis provides the most direct and rapid way to make oligonucleotides. During standard solid-phase synthesis, the growing oligonucleotide is anchored to a Controlled Pore Glass (CPG) silica support *via* a succinamide linker.[65] Removal of the oligonucleotide from the CPG support requires treatment with concentrated ammonia for several hours or for shorter times at elevated temperatures.[65,138] Such basic conditions are incompatible with oligonucleotides containing base-sensitive ester and amide functions. To date, the synthesis of oligonucleotides with base-sensitive functions has relied upon solution chemistry which can be time consuming, indirect, low yielding and therefore impractical for synthesis of multiple sequence analogues for biological evaluation.

The following points were considered in devising a new strategy for delivery of oligonucleotides into cells:

i. O-oligos are achiral and have the greatest ability to induce RNase-H activity and have the highest affinity for target RNA.

ii. Acyloxymethyl protection of phosphorus offers possible protection against nuclease enzymes.

iii. Intracellular carboxyesterases have been shown to hydrolyse acyloxymethyl protected S-oligos yielding the parent S-oligo.

In addition, solution chemistry which has been used to date for synthesis of acyloxymethyl-protected oligonucleotides is impractical for long sequences. Furthermore, the conditions required to deprotect the conventional succinamide linker used in standard solid-phase synthesis is incompatible with acyloxymethyl phosphate protection.

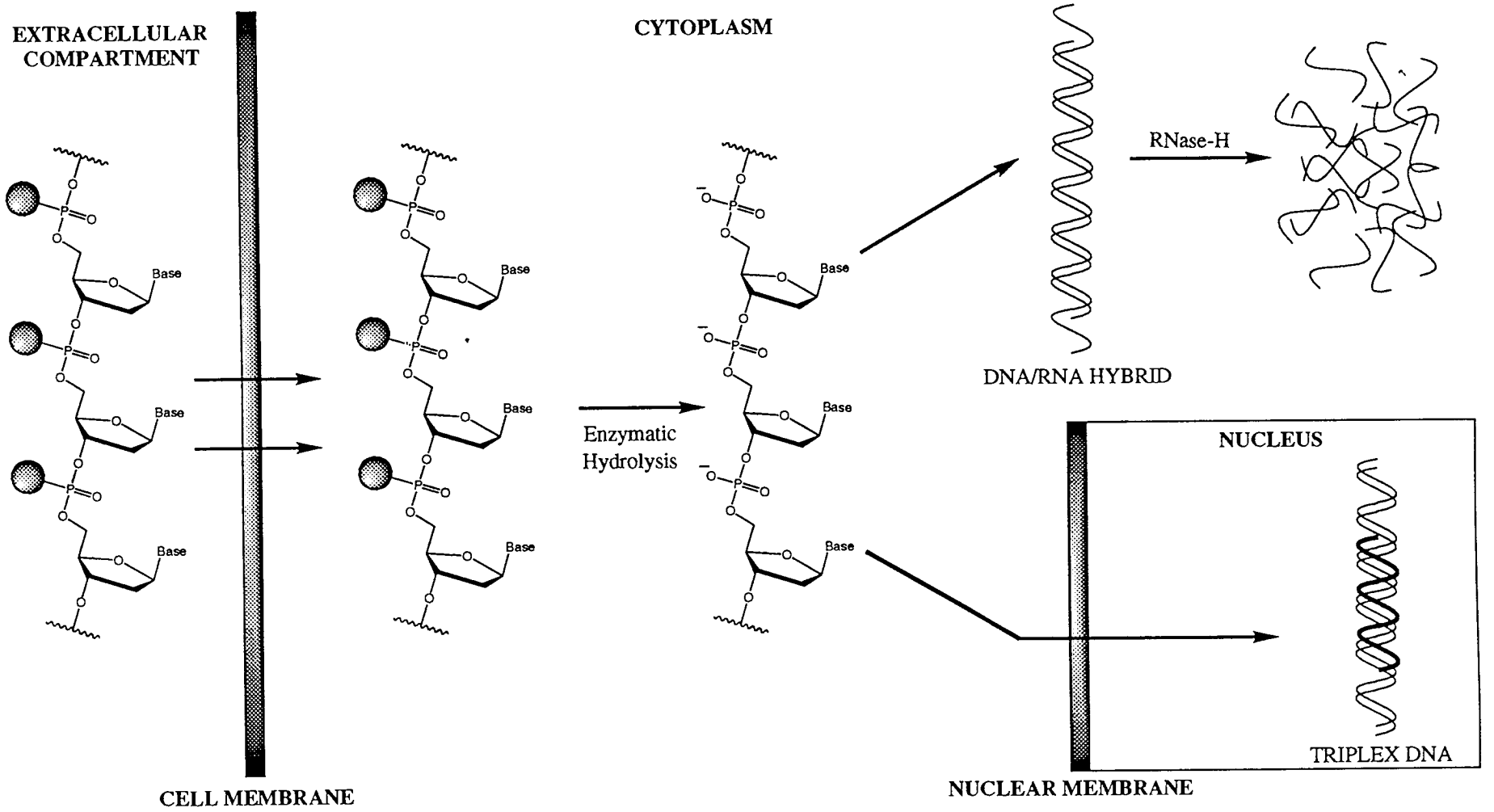


Figure 5.4 Delivery of prooligonucleotides into cells.

A novel linker between CPG and the oligonucleotide, cleavable under neutral conditions, is therefore required to allow automated solid phase synthesis of oligonucleotides sensitive to aqueous ammonia. The use of a photolabile linker has been reported but the formation of thymine-thymine dimers has proved to be a major drawback.[192,193] A silyl based linkage between CPG and a base-sensitive oligonucleotide could potentially yield the intact base-sensitive oligonucleotide on fluoride deprotection.

5.4 Design of a Silyl-Linked Support for Solid Phase DNA Synthesis

Silyl protecting groups have been used extensively in nucleoside chemistry to mask amines,[194,195] alcohol's[196] and phosphorus[197] functions. The 2'-OH function of ribonucleotides require protection during automated solid phase synthesis of RNA.[65,138] One of the most successful ribose protection strategies for the 2'-OH of oligonucleotides uses the fluoride labile *tert*-butyldimethylsilyl group for automated RNA synthesis. Danishefsky *et al* (1993)[198] developed a solid phase synthesis method to assemble oligosaccharides on a diphenylsilyl-linked polystyrene support. Cleavage of the oligosaccharide from the solid support was achieved during 4 hours using tetra-*n*-butylammonium fluoride (TBAF) buffered with acetic acid.

It was reasoned that a diisopropylsilyl-based linker would be beneficial in this project to anchor a 4,4'-dimethoxytritylated nucleoside to CPG for solid phase synthesis of oligonucleotides. It was necessary therefore to test whether the linkage was:

i. stable during automated oligonucleotide solid phase synthesis, and

ii. cleavable giving the oligonucleotide product using mild fluoride deprotection conditions.

5.5 Chemical Synthesis

Dr. Anne Routledge, a colleague in the Aston group, carried out the initial chemical synthesis of the novel diisopropyldisiloxane linker between a nucleoside and CPG,

illustrated in Figure 5.5. All subsequent analyses and cleavage/stability studies of this material were carried out as part of this present work.[199]

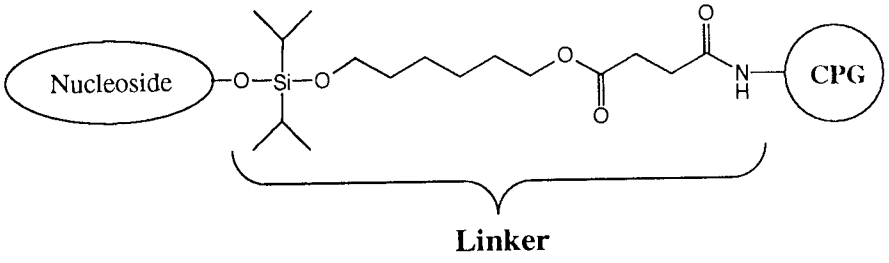


Figure 5.5 Novel diisopropyldisiloxane linker between a nucleoside and CPG.

5'-*O*-(4,4'-Dimethoxytrityl)-2'-deoxythymine (DMTr-dT) was selected for attachment to CPG since no additional protection of the base was required and triflate **106** was already known in the literature.[200,201] Succinate **101** was prepared by monobenzylation of 1,6-hexanediol, followed by reaction of the protected alcohol with succinic anhydride. Pentafluorophenyl ester **103** was synthesised in two steps from succinate **101** *via* **102** as shown in Figure 4.6. 2'-Deoxythymine **104** with its electrophilic 3'-O-diisopropylsilyl triflate underwent nucleophilic attack by spacer group **103** giving the silyl-linked nucleoside adduct **105**. Commercially available CPG silica was derivatised with silyl-linked adduct **105** giving the solid support material **106**.

5.6 Evaluation of the Silyl-Linked support 106 for Automated Synthesis

The experimental work of this chapter is an investigation of the viability uses of the diisopropylsilyl based linker in solid phase synthesis. The following aspects were studied:

i. stability of the silyl linkage during automated solid phase synthesis;

ii. optimisation of deprotection conditions of the oligonucleotide using fluoride and

iii. the viability of this technology for synthesis of base-sensitive oligonucleotides.

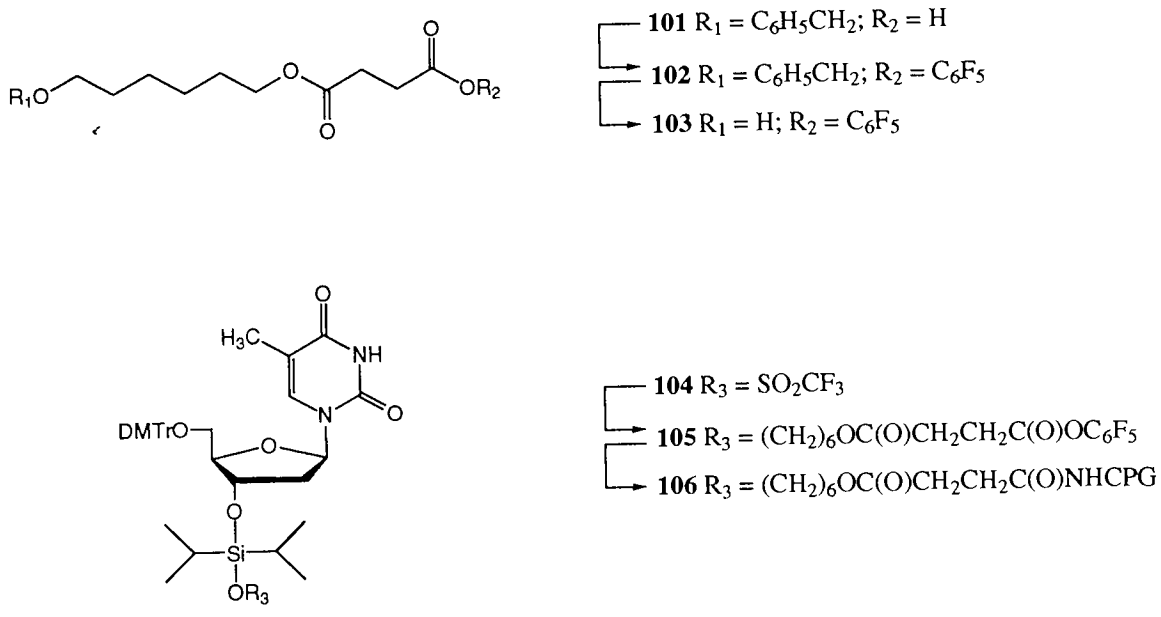


Figure 5.6 Synthesis of the novel silyl-linked solid support **106**.

5.6.1 Determination CPG solid support loading

The efficiency of the coupling reaction between **105** and the CPG support was determined by accessing the loading through assaying the release of $DMTr^+$ cation following treatment of **106** with a mild acid (Figure 5.7).

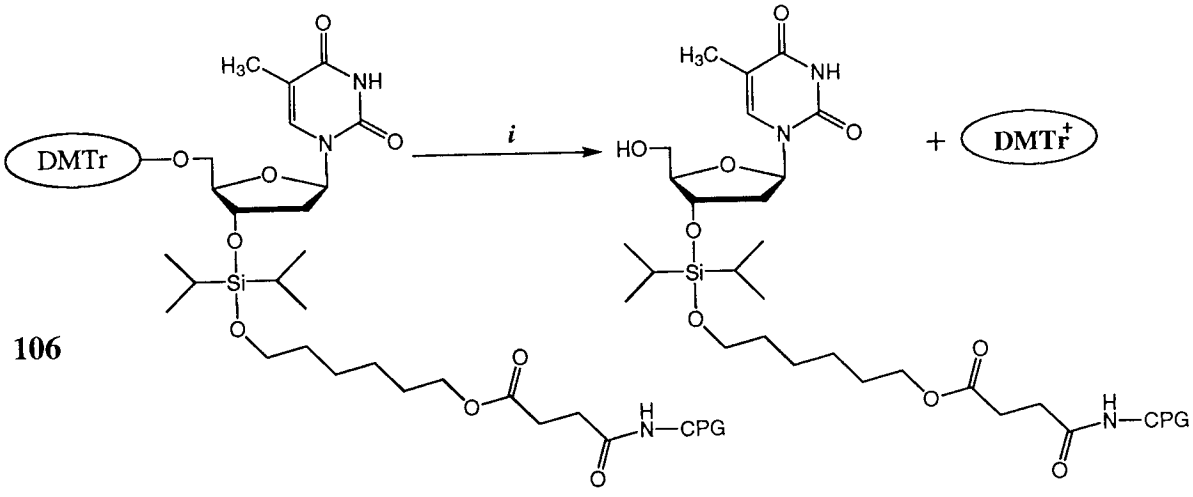


Figure 5.7 Determination of CPG loading. Reagent *i* 70% perchloric acid/MeOH.

Approximately 2 mg of **106** was accurately measured and treated with of 70% perchloric acid in methanol (2 mL) and the solid support was filtered off using a 4 μm pore PTFE syringe filter. The acid cleaves 4,4'-dimethoxytrityl from the 5'-position of thymine bound to the CPG solid support. To ensure full recovery of the $DMTr^+$ cation, the filtered CPG solid support was washed three times with 0.75 mL of 70% perchloric acid in methanol. The combined DMTr cation solutions were accurately diluted to 5 mL with 70% perchloric acid in methanol in a volumetric flask (5 mL). The absorbance of the orange coloured solution was measured at 498 nm using a UV cell of 1 cm path length. The deprotection experiment was performed three times and the absorbance calculated using Equation 5.1.[65] An averaging of the results showed that a loading of 27 μmol g^{-1} had been achieved.

$$\text{Loading} = \frac{\text{Abs. at } 498 \times \text{vol. (mL) of } HClO_4 \times 14.3}{\text{wt. of support (mg)}} \qquad \text{[Eqn. 5.1]}$$

5.6.2 Deprotection of the silyl-linked solid support

The susceptibility of the diisopropylsilyl linker to attack by fluoride to liberate 2'-deoxythymidine was examined. Approximately 2 mg of **106** was accurately weighed and treated with 1 M TBAF in THF (0.4 mL) for 60 seconds at room temperature. The solid support was filtered off through a 4 μm pore PTFE syringe filter and washed with methanol (3 x 0.4 mL). The filtrate containing the cleaved 5'-*O*-(4,4'-dimethoxytrityl)-2'-deoxythymidine was accurately diluted to 5 mL with 70% perchloric acid in methanol and its absorbance measured at 489 nm. 4,4'-Dimethoxytrityl cation concentration in this solution correlates directly with the amount of 2'-deoxythymidine cleaved from the solid support using fluoride deprotection (Figure 5.8). Equation 4.1 allows a comparison between loading and cleavage of 2'-deoxythymidine from the solid support. The 4,4'-dimethoxytrityl assay indicated that quantitative cleavage of 2'-deoxythymidine from the solid support had been achieved during 60 seconds at room temperature.

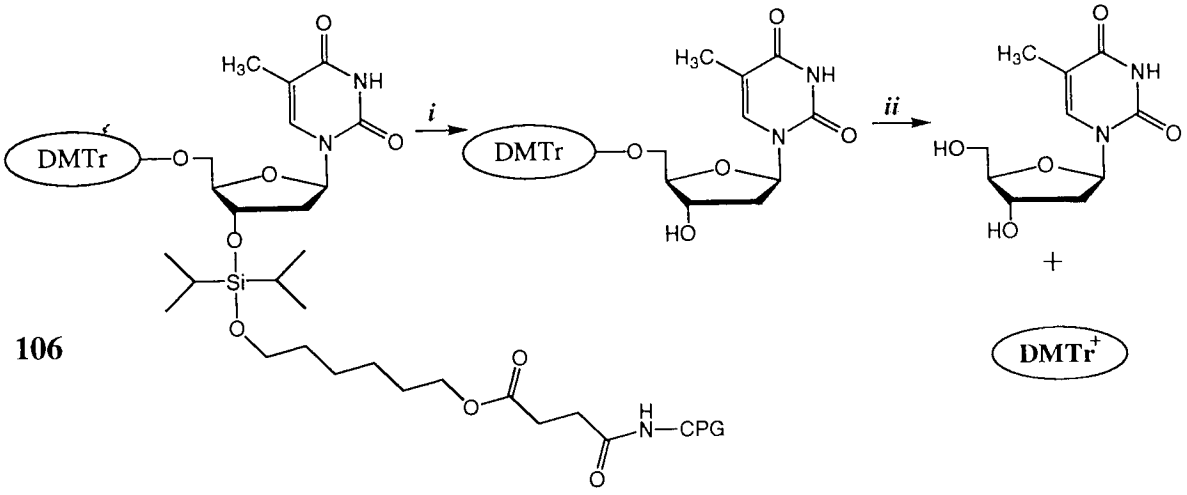


Figure 5.8 Determination of cleavage efficiency of the diisopropyl linker **106**. Reagents *i* 1 M TBAF; *ii* 70% perchloric acid in methanol.

It was necessary to find the mildest possible deprotection conditions to facilitate rapid and quantitative cleavage of the nucleoside from the solid support. Four parameters were considered:

i. concentration of TBAF;

ii. cleavage time;

iii. reaction temperature and

iv. basicity of the deprotection solution.

Deprotection of dT from CPG was conducted using 0.5 M TBAF in THF for 60 seconds at room temperature. Quantitative recovery of dT was again achieved as indicated by 4,4'-dimethoxytrityl cation assays. Satisfied with the serendipitous optimisation of parameters **i-iii**, the pH of the TBAF solution was considered.

Deprotection of dT from the diisopropylsilyl linker is thought to be facilitated by nucleophilic attack on silicon by fluoride. However, commercially available 1M TBAF in THF solution contains 5% water and is basic as shown by universal indicator paper. The pH of the deprotection solution was reduced by buffering with one equivalent of acetic acid

before cleaving silyl-linker dT **106** from CPG at varying concentrations of fluoride ion (Table 5.2).

TBAF concentration/M	1	0.5	0.1
Acetic acid concentration/M	-	0.5	0.1
% Recovery of dT	100	100	70

Table 5.2 Deprotection conditions for cleaving **106** using TBAF solutions.

It was found that using 0.5 M TBAF and 0.5 M acetic acid in THF (0.4 mL) successfully cleaved dT quantitatively from the silyl-linked solid support **106** (10 mg) within 60 seconds at room temperature. Concentration of TBAF was considered to be the controlling factor for linker cleavage since the volume of deprotection solution was kept constant in each case.

5.6.3 Synthesis and deprotection of d(Tp)$_7$T using silyl linker 106.

Once the ease with which dT could be removed from the solid support had been demonstrated, it was then necessary to determine the stability of the silyl linkage during automated DNA synthesis conditions. The sequence d(Tp)$_7$T was synthesised using the silyl-linked solid support **106**. Commercially available cyanoethyl protected 2'-deoxythymidine phosphoramidite was used. The heterocycle of dT requires no protection for automated DNA synthesis thus avoiding additional deprotection steps. Although acyloxymethyl-protected phosphoramidites are the ultimate synthetic aim for the pro-drug approach outlined above, the use of commercially available cyanoethanol protected phosphoramidites demonstrated the viability of the linker for automated oligonucleotide synthesis.

Each addition of a new nucleoside phosphoramidite to the growing oligonucleotide during automated synthesis is preceded by acidic cleavage of 4,4'-dimethoxytrityl group of the C5'

protected hydroxyl from the oligonucleotide. The quantity of cleaved 4,4'-dimethoxytrityl cation can be assayed by UV spectrophotometry giving a measure of the synthesis yield (Chapter 3). Any significant reduction in 4,4'-dimethoxytrityl concentration may indicate:

i. poor coupling between the solid-supported oligonucleotide and the incoming nucleoside phosphoramidite, often due to moisture or out-of-date reagents;
ii. chemical instability of the linker to the reagents used during the automated oligonucleotide synthesis cycles.

Oligonucleotide d(Tp)$_7$T was made using a standard succinamide linker on a 0.2 μmol scale, monitored by 4,4'-dimethoxytrityl cation assay (Figure 5.9, (a)), which indicated that the DNA synthesiser and the reagents used were in good working order. The silyl-linked support **106** was use to prepare d(Tp)$_7$T on a 0.7 μmol scale (Figure 5.9, (b)).

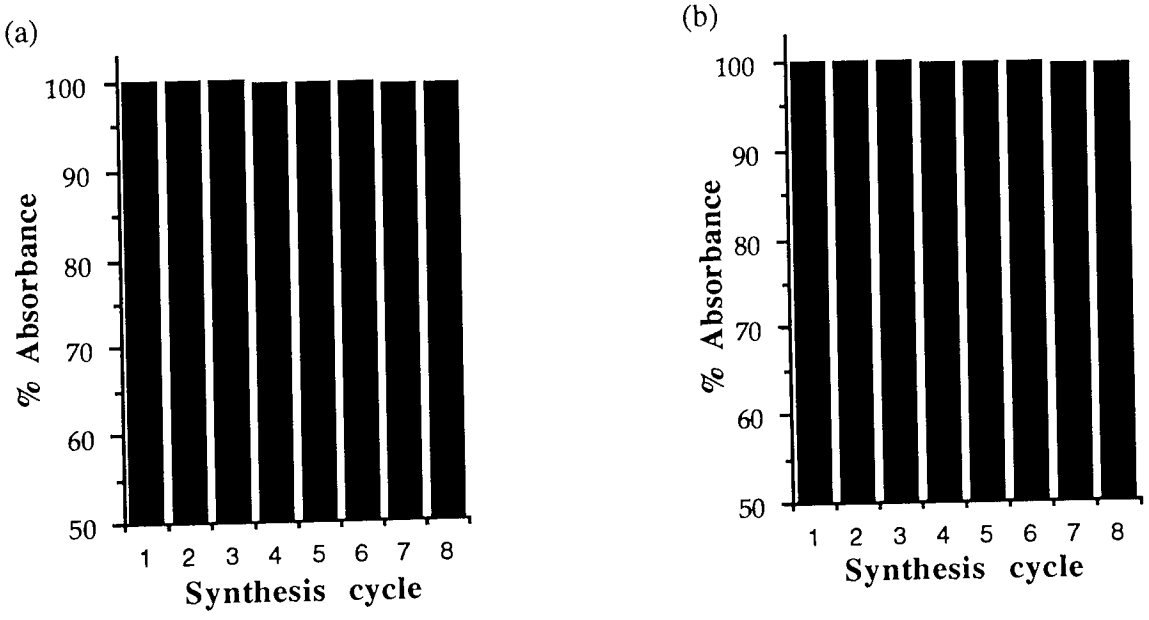


Figure 5.9 4,4'-Dimethoxytrityl cation assays for the synthesis of (a) d(Tp)$_7$T using a standard succinamide linker and (b) silyl-linked support **106**.

The collected 4,4'-dimethoxytrityl fractions were diluted with of 1.0 M 4-toluenesulphonic acid in acetonitrile (10 mL), and accurately diluted to 20 mL with acetonitrile and their absorbance determined at 498 nm. Absorbance values for 4,4'-dimethoxytrityl fractions for each coupling were compared with the second 4,4'-dimethoxytrityl fraction as a percentage. The first fraction can give a misleading high result due to detritylation of the support bound nucleoside during storage. The overall yields for the syntheses of $d(Tp)_7T$ conducted using the succinamide and silyl linkers, were 98.9% and 99.2% respectively.

Once the chemical integrity of the silyl linker had been established it was necessary to cleave the oligonucleotide from the CPG support and analyse the product. The synthesised oligonucleotide attached to the solid support has cyanoethyl-protected phosphate centres. Standard deprotection conditions using concentrated aqueous ammonia serves to cleave the oligonucleotide from the succinamide linker and unmask the cyanoethanol protected phosphate centres by β-elimination. Deprotection of dT from the silyl-linked support had been shown to occur practically instantaneously using TBAF. It was reasoned that fluoride could potentially act as a base, facilitating β-elimination of the cyanoethyl protected phosphate centres (Figure 5.10).

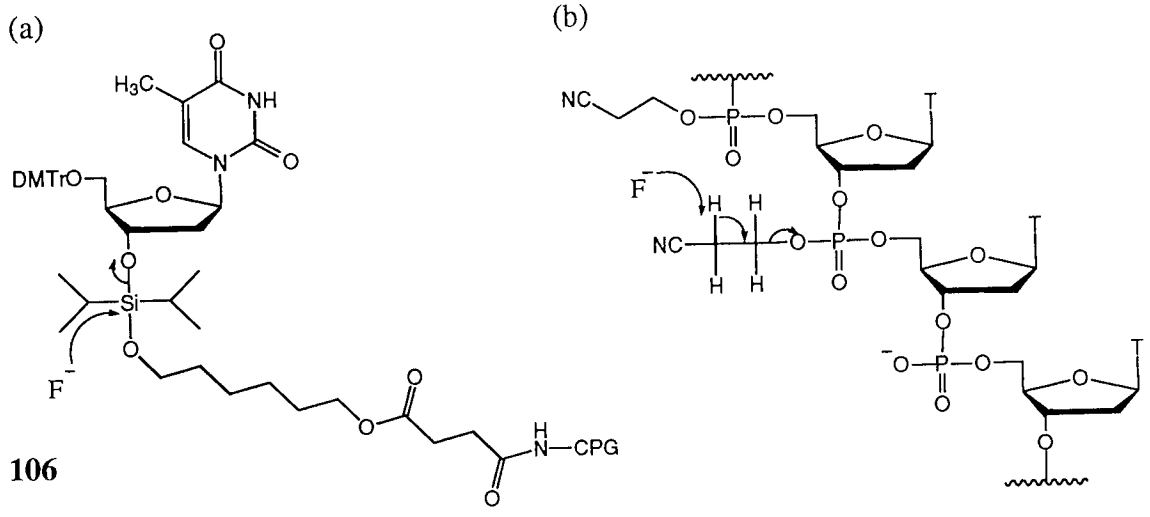


Figure 5.10 Deprotection of (a) **106** and (b) cyanoethyl-protected phosphate centres of $d(Tp)_7T$ with F^-.

Deprotection of $d(Tp)_7T$, synthesised using **106**, was monitored by HPLC. Authentic $d(Tp)_7T$ synthesised using conventional methods was used as a reference sample. Although cleavage of the silyl linkage was virtually instantaneous, heating **106** with 0.5 M TBAF (0.4 mL) for 2 hours at 50°C was necessary to deprotect the phosphorus centres. A dark brown colour developed during deprotection indicating possible degradation of the CPG solid support. Crude samples of the deprotection solution (50 μL) were analysed by gradient HPLC (Section 6.2). The profiles of $d(Tp)_7T$ synthesised using a standard succinamide linker (a), and silyl linker (b) are shown (Figure 5.11). Profile (c) shows (a) and (b) co-elute, indicating the two samples to be identical. Using TBAF buffered with acetic acid (Table 5.2) produced comparable results by HPLC but no brown colour developed during cleavage/deprotection. This implies that the addition of acetic acid offers protection against chemical degradation of the solid support which is itself composed of Si-O bonds.

It was necessary to unambiguously establish the structure of the oligonucleotide synthesised using the diisopropylsilyl linker. Time-of-flight mass spectrometry (MALDI-TOF, Chapter 3) can provide a value for the molecular ion peak, equivalent to molecular weight. Controlled pore glass contains silicon-oxygen bonds which are susceptible to attack by TBAF, leading to degradation of the solid support which might adversely affect MALDI-TOF analysis. A SEP-PAK column was used to remove CPG (both low and high molecular weight) and TBAF from the oligonucleotide sample prior to time-of-flight mass spectrometry. The packed silica of the SEP-PAK column acted in two ways:

i. as a filter to remove particulate CPG, and

ii. as a reverse-phase chromatography column to retain the less hydrophilic oligonucleotide allowing the more hydrophilic TBAF to be washed away.

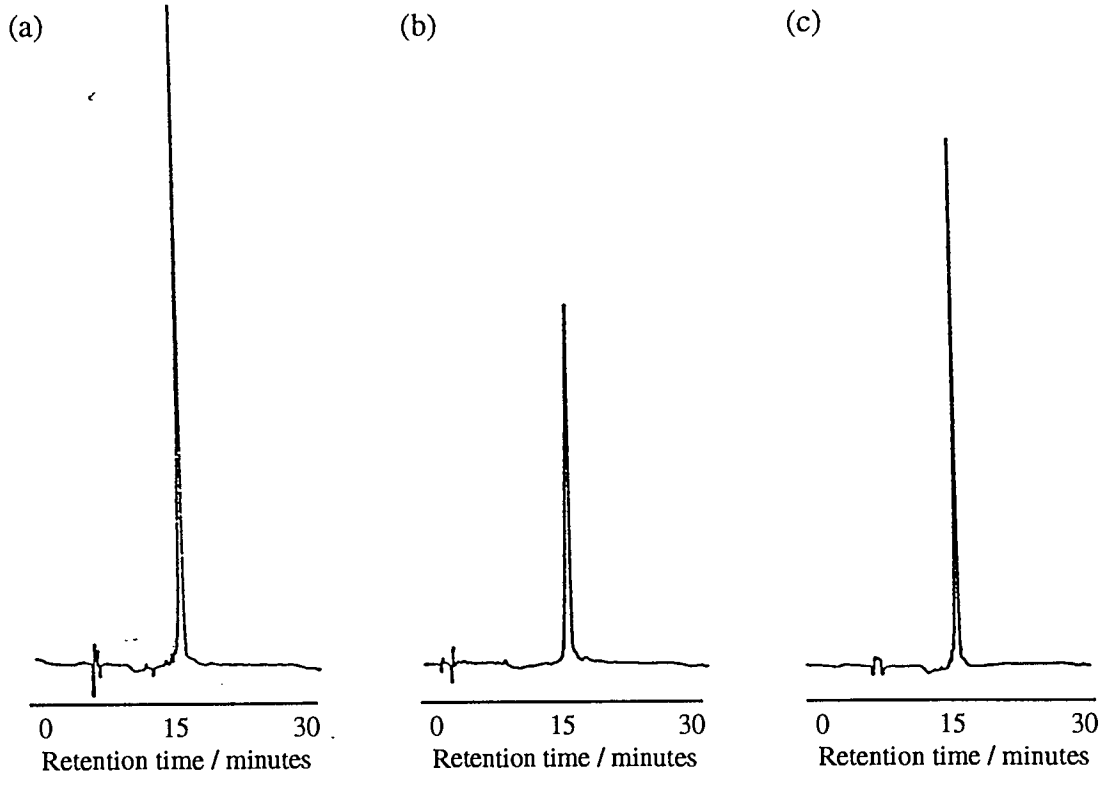


Figure 5.11 HPLC elution profiles of $d(Tp)_7T$ made using a standard succinamide linker (a) and the diisopropylsilyl linker (b). Profile (c) shows a co-injection of $d(Tp)_7T$ made using both solid supports.

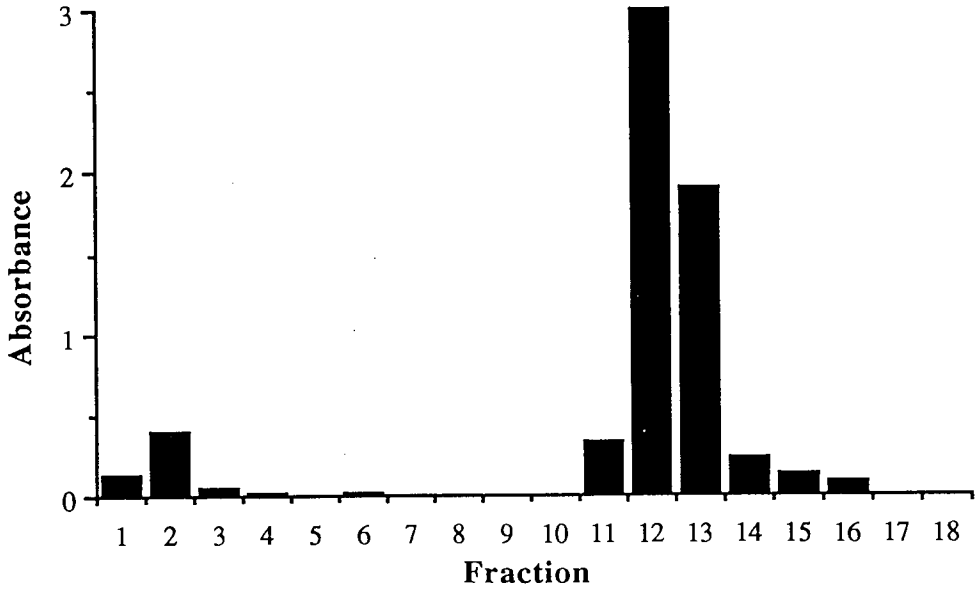


Figure 5.12 UV profile of SEP-PAK purification of $d(Tp)_7T$.

The TBAF-deprotected oligonucleotides were purified in the following way:

i. The SEP-PAK column was conditioned by passing through it 50 mL of HPLC grade methanol followed by 200 mL of HPLC grade water.
ii. The crude sample mixture (approx. 0.4 mL) was loaded onto SEP-PAK column and eluted with 1.0 mL of HPLC grade water.
iii. The filtrate was collected in a 1.5 mL UV cuvette with a 1 cm path length of.
iv. The absorbance of the sample was determined at 260 nm and the sample transferred to a 1.5 mL Ependorf tube.
v. Water (1.4 mL) was passed through the SEP-PAK column and the filtrate collected in the UV cuvette for absorbance determination. This process was repeated for 9 cycles.
vi. On the eleventh cycle 50% HPLC acetonitrile in water was used to collect the oligonucleotide in 1.5 mL aliquots as described above for 4 cycles.

The UV profile for this purification process is shown (Figure 5.12). The absorbance of fraction 12 is off-scale. Fractions 2 and 10-15 were concentrated under vacuum. Fraction 2 was suspended in D_2O and analysed by ^{19}F NMR, referenced upfield of CF_3CCl_3. A single peak at -43.6 was observed (Figure 5.13) indicating the presence of fluoride.

Fraction 10 was analysed by ^{19}F NMR but no peak was observed after 800 scans indicating that no significant quantities of F^- were present. Fractions 11 to 14 were suspended in HPLC water, combined, and analysed by HPLC. The sample was shown to co-elute with authentic $d(Tp)_7T$ indicating stability of the oligonucleotide during the SEP-PAK de-salting procedure. The sample was concentrated under vacuum and subjected to MALDI-TOF mass spectral analysis. The major peak (100%) was observed at 2369 (+/- 2) m/z, compared with the calculated molecular mass of 2370.

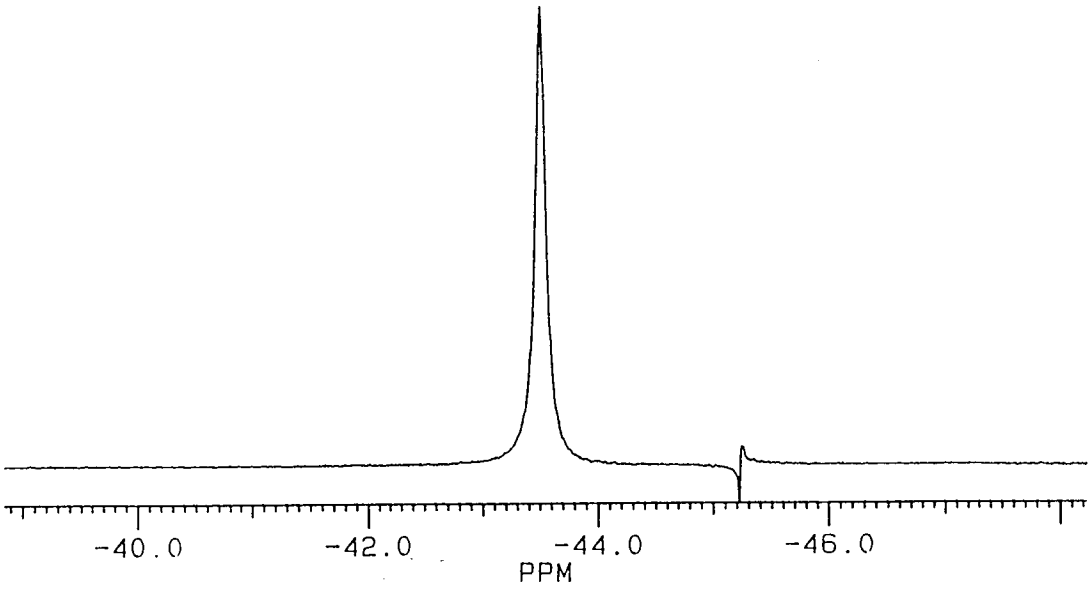


Figure 5.13 ^{19}F NMR spectrum of fraction 2 from SEP-PAK purification of $d(Tp)_7T$.

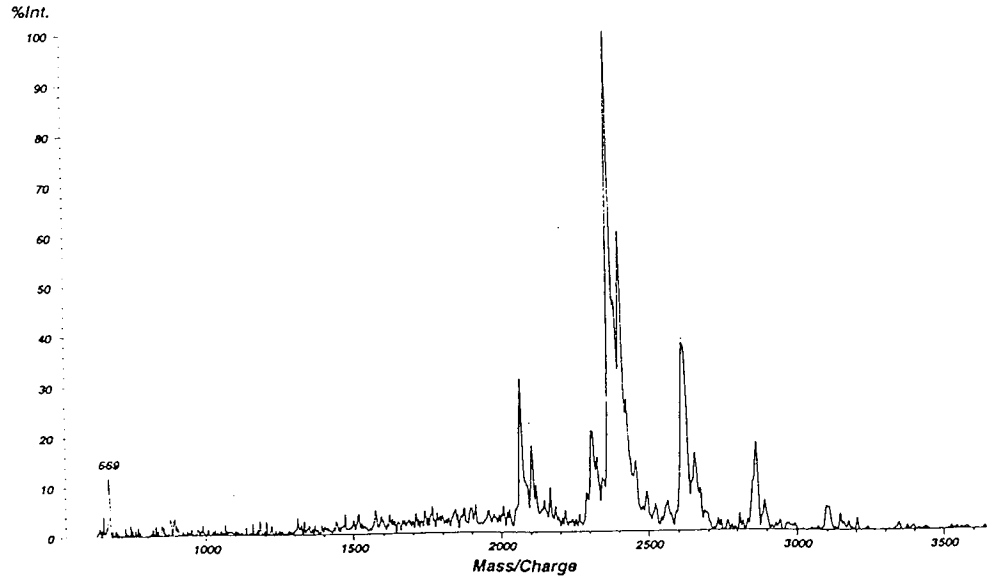


Figure 5.14 MALDI-TOF spectrum of $d(Tp)_7T$.

5.6.4 Synthesis and deprotection of d(Tps)$_7$T using the silyl-linked solid support 106

The automated solid phase synthesis of phosphorothioate oligonucleotides is essentially identical to that of phosphodiester oligonucleotides except for the oxidation step. The 3'-P(III) centre of the newly attached nucleoside is either oxidised by I_2/THF/H_2O solution to make phosphodiester oligonucleotides or oxidised, in this work, by tetraethylthiuram disulphide in acetonitrile to make a phosphorothioate oligonucleotides.

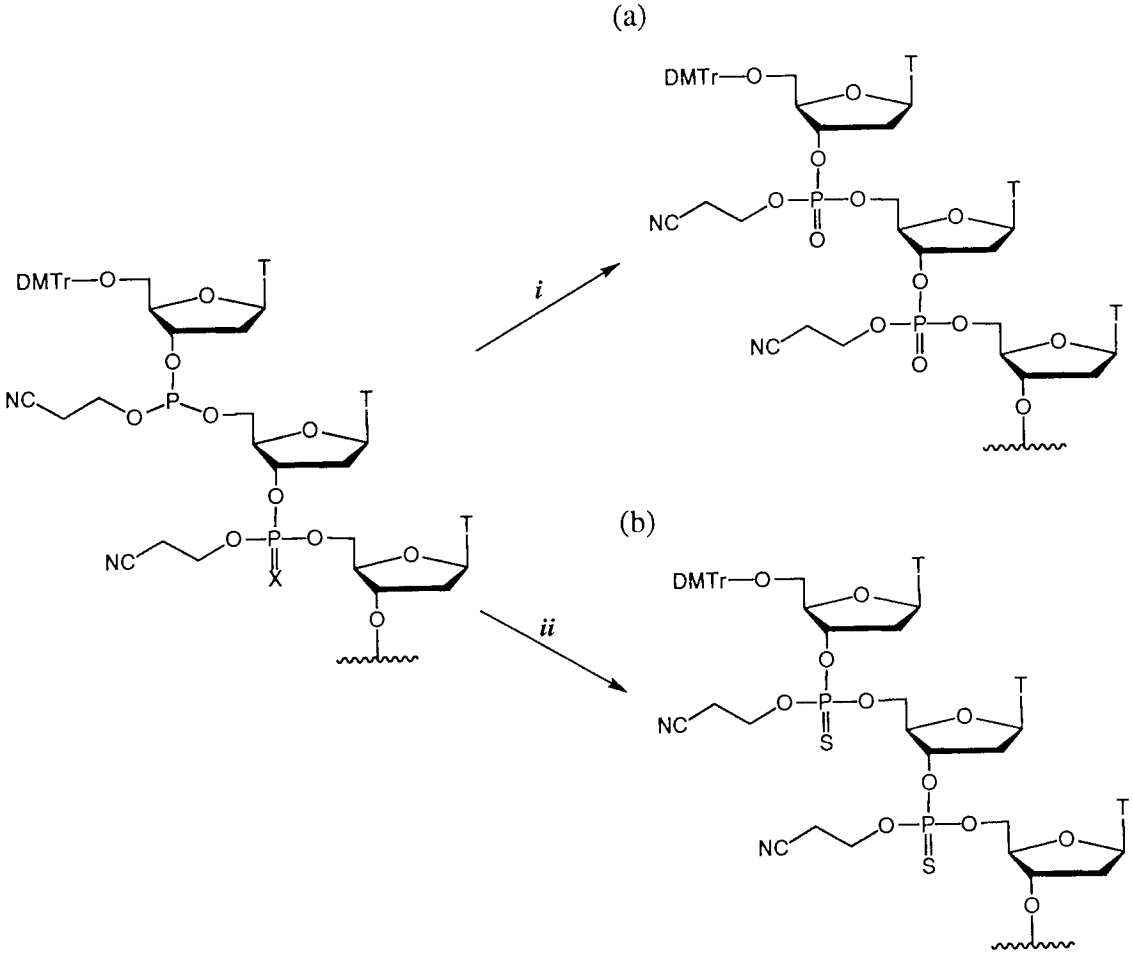


Figure 5.15 Phosphoramidite oxidation to (a) phosphodiester or (b) phosphorothioate nucleosides; X = O or S; Reagents: *i* I_2/THF/H_2O; *ii* tetraethylthiuram disulphide/CH_3CN.

The chemically stability of the diisopropyl linkage to the oxidising solution, I_2/THF/H_2O, has been demonstrated during synthesis of the phosphodiester oligonucleotide d(Tp)$_7$T.

The chemically stability of the diisopropyl linkage towards the sulphurising agent tetraethylthiuram disulphide was then examined. The oligonucleotide d(Tps)$_7$T was synthesised using standard conditions with cyanoethyl phosphoramidites. Oxidation was achieved using tetraethylthiuram disulphide. Assay of cleaved 4,4'-dimethoxytrityl showed an overall coupling efficiency of 99.1%. Deprotection of d(Tps)$_7$T from CPG was achieved using 1M TBAF in THF at 50°C during 2 hours. Figure 5.16 (a) shows the HPLC profile corresponding to a crude sample of the phosphorothioate d(Tps)$_7$T made using a standard succinamide-linked CPG silica solid support whereas (b) shows the crude d(Tps)$_7$T sequence made using the silyl linker. Co-injection of a mixture of d(Tps)$_7$T made using both linkers gave a single peak (c), showing the deprotected oligonucleotides to be identical.

5.6.5 Synthesis and deprotection of a base-labile oligonucleotide

In order to determine whether solid support **106** in combination with the TBAF deprotection strategy could yield oligonucleotides containing base-sensitive functional groups, the phosphodiester d(Tp)$_3$(C^{Bz}p)$_2$(Tp)$_2$T **108** octomer was synthesised. Benzamide (Bz) protected C phosphoramidites were used as an example of a base-sensitive nucleoside. The amino protected benzamide function is readily cleaved by to aqueous ammonia solution. Isolation of the oligonucleotide d(Tp)$_3$(C^{Bz}p)$_2$(Tp)$_2$T would demonstrate the potential application of **106** with TBAF deprotection for synthesis of base-sensitive oligonucleotides. Automated oligonucleotide synthesis was carried out using 10 mg of **106** giving the crude d(Tp)$_3$(C^{Bz}p)$_2$(Tp)$_2$T in 99% yield as determined by 4,4'-dimethoxytrityl assay. Cleavage of d(Tp)$_3$(C^{Bz}p)$_2$(Tp)$_2$T from the silyl-linked support was achieved within seconds using 0.5M TBAF and 0.5M acetic acid in THF, although removal of the cyanoethyl protecting groups required heating at 50°C for 2 hours. The HPLC profile of a crude sample of d(Tp)$_3$(Cp)$_2$(Tp)$_2$T **107** made using a standard succinamide-linked CPG silica solid support, cleaved with concentrated aqueous ammonia at 55°C for 16 hours, is shown (Figure 5.17 (a)). Under the ammonia deprotection conditions, benzamide-protected amino groups are unmasked. The HPLC profile of the crude benzamide-protected sequence

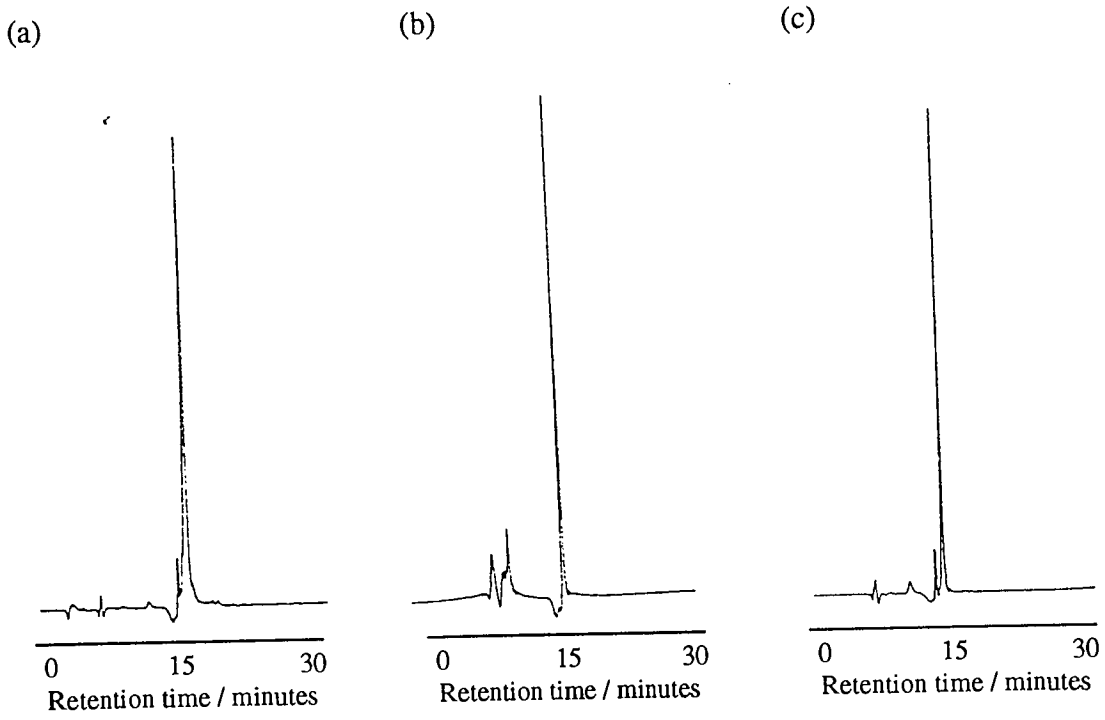


Figure 5.16 HPLC profile of (a) a standard succinamide linker; (b) d(Tps)$_7$T made using the support bound silyl linker **106** and (c) co-injection (a) and (b).

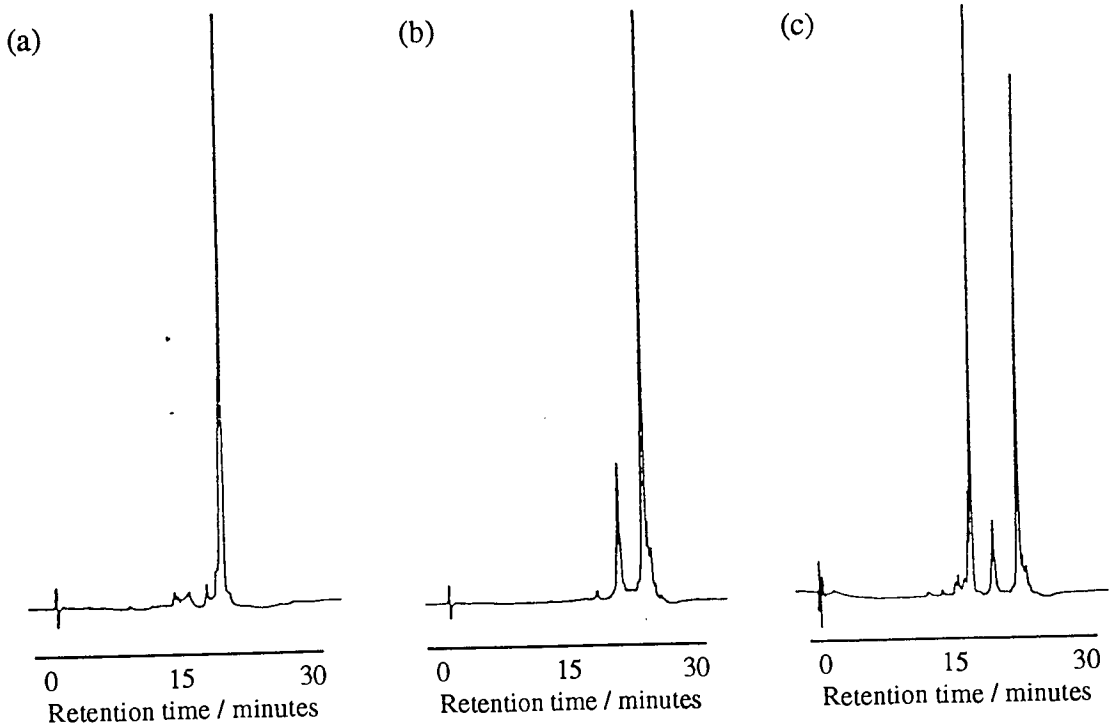


Figure 5.17 HPLC profile of (a) oligonucleotide **107** made using a standard succinamide linker; (b) oligonucleotide **108** made using the CPG support bound silyl linker **106** and (c) co-injection of (a) and (b).

$d(Tp)_3(C^{Bz}p)_2(Tp)_2T$ made using the silyl-linked support **106** is shown in HPLC profile (b). Figure 5.17 (c) shows the HPLC profile of a co-injection of samples $d(Tp)_3(Cp)_2(Tp)_2T$ and $d(Tp)_3(C^{Bz}p)_2(Tp)_2T$. Two major peaks were noted, $d(Tp)_3(Cp)_2(Tp)_2T$ and a peak with a longer retention time, due to $d(Tp)_3(C^{Bz}p)_2(Tp)_2T$.

It was important to unambiguously establish that the benzamide protecting groups in $d(Tp)_3(C^{Bz}p)_2(Tp)_2T$ were intact after fluoride deprotection of the oligonucleotide. The crude sequence was separated from the solid support and desalted using a SEP-PAK column as described. $d(Tp)_3(C^{Bz}p)_2(Tp)_2T$ was then purified by HPLC and subjected to MALDI-TOF mass spectrometric analysis (Figure 5.18). A molecular peak was observed at 2548 m/z which was identical to the molecular weight calculated for $d(Tp)_3(C^{Bz}p)_2(Tp)_2T$ corresponding to the molecular formula $C_{92}H_{111}N_{18}O_{54}P_7$.

The complementary sequence to oligonucleotide **108**, $d(Ap)_3(Gp)_2(Ap)_2T$ (**109**), was synthesised using the standard succinamide linker and phosphoramidite chemistry. The affect of the amino protected benzamide function on duplex stability was assessed using thermal analysis (Chapter 3)

5.5 Conclusion

The deprotection strategy described, using the diisopropylsilyl linker is direct, efficient and compatible with base-sensitive benzamide substituents. This methodology has the potential to be extended to allow the synthesis of base-labile acyloxymethyl phosphate protected oligonucleotides exploring their use as antisense applications. Work using the silyl-linked support **106** has continued within the Aston group, where the technology is being used to synthesise an oligonucleotide conjugate of the base-sensitive anti-cancer imidazotetrazinone mitozolomide.[203]

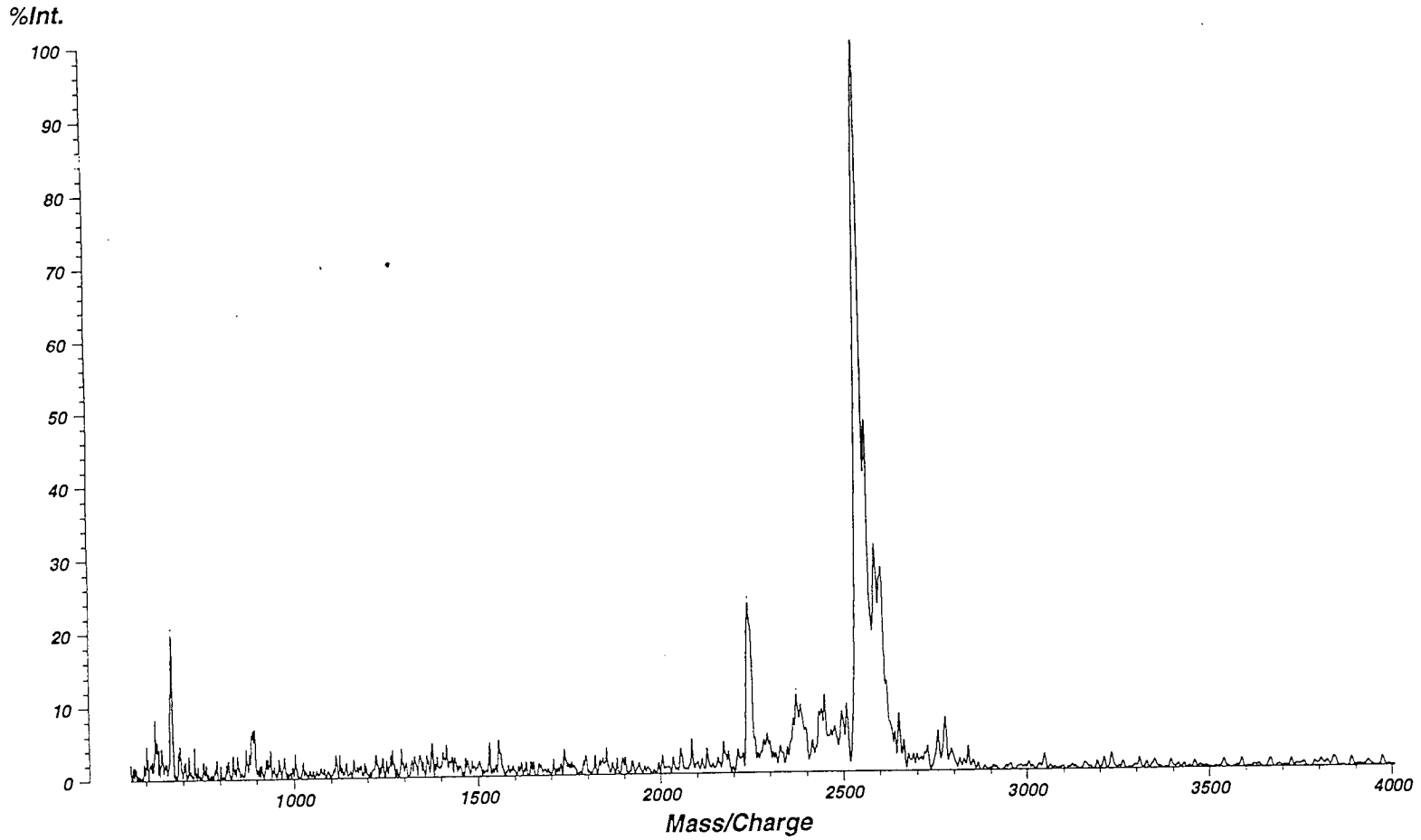


Figure 5.18 MALDI-TOF spectrum of $d(Tp)_3(C^{Bz}p)_2(Tp)_2T$ **108.**

Chapter 6

Experimental

6.1 Chemistry

6.1.1 General methods

NMR spectra were recorded on an a AC250 spectrometer at ^{1}H (250.1 MHz), ^{13}C (62.9 MHz) and ^{19}F (235.3 MHz). Positive chemical shifts are downfield of tetramethylsilane for ^{1}H and ^{13}C and negative chemical shifts are upfield of CF_3CCl_3 reference for ^{19}F. Where necessary, proton and carbon assignments were made from the results of NOE and ^{13}C-H heteronuclear correlation experiments. 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 $LSIMS^+$ mode using a VG ZabSpec instrument. Infrared spectra were recorded on a Mattson Galaxy 2020 FT-IR Spectrophotometer. Ultraviolet spectra were recorded using Varian Carey 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).

6.1.2 Synthetic chemistry

1-(3',5'-Di-O-acetyl-2'-deoxy-β-D-erythro-pentofuranosyl))-5-methyl-pyrimidine-2,4(1H,3H)-dione 1: 2'-Deoxythymidine (1 g, 4.13 mmol) was dried by co-evaporation with anhydrous pyridine (3x15 mL), dissolved in pyridine (40 mL) and placed under Ar. A catalytic amount of dimethylaminopyridine (30 mg) was added and the solution stirred at 0°C for 15 minutes. Acetic anhydride (4.57 mL, 41.3 mmol) was added and the solution

stirred at 4°C for 14 h before being concentrated under vacuum. Flash chromatography, eluting with EtOAc, yielded the title compound as a foam (1.14 g, 85%). The analytical sample was obtained by crystallising from EtOAc; TLC (EtOAc): R_f 0.34; ^{1}H NMR [$(CD_3)_2SO$]:δ 1.78 (s, 3 H, 5-CH_3), 1.98 (s, 3 H, acetyl CH_3), 2.05 (s, 3 H, acetyl CH_3), 2.27 (m, 1 H, 2'-CH), 2.43 (m, 1 H, 2'-CH), 4.15 (m, 3 H, 4'-CH, 5'-CH), 5.17 (m, 1 H, 3'-CH), 6.17 (t, 1 H, *J* 6.2 Hz, 1'-CH), 7.49 (s, 1H, 6-CH), 11.40 (s, 1H, NH).

1-(3',5'-Di-O-acetyl-2'-deoxy-β-D-erythro-pentofuranosyl)-5-methyl-4-(1,2,4-triazolyl)-pyrimidine-2(1H)-one 2: Nucleoside **1** (1.00 g, 3.09 mmol) was dried by co-evaporation with anhydrous pyridine (3x30 mL), dissolved in pyridine (50 mL), placed under Ar and the solution cooled to 0°C. 4-Chlorophenyl phosphorodichloridate (1.00 mL, 6.17 mmol) was added and the mixture was stirred for 5 minutes. 1,2,4-1*H*-Triazole (0.85 g, 12.3 mmol) was added and the mixture was stirred at room temperature for a further 72 h before being concentrated under vacuum. Flash chromatography, eluting with EtOAc, yielded the title compound (0.55 g, 49%) as a foam; m.p. 127°C; TLC (EtOAc): R_f 0.36; UV (95% EtOH) λ_{max} 252 nm (13800), 238 nm (9850); IR (KBr disc): υ_{max} 3465, 3104, 2933, 2902, 1745, 1679, 1525, 1513, 1425, 1380, 1238, 1068, 978, 775, 669 cm^{-1}; ^{1}H NMR [$(CD_3)_2SO$]: δ 2.03 (s, 3 H, CH_3 (acetyl)), 2.08 (s, 3 H, CH_3 (acetyl)), 2.32 (s, 3 H, 5-CH_3) , 2.49 (m, 2 H, 2'-CH_2), 4.35 (m, 3 H, 4'-CH, 5'-CH_2), 5.22 (m, 1 H, 3'-CH), 6.15 (t, 1 H, *J* 6.5 Hz, 1'-CH), 8.25 (s, 1 H, 6-CH), 8.39 (s, 1 H, 3-CH (triazolyl)), 9.33 (s, 1 H, 5-CH (triazolyl)); ^{13}C NMR [$CDCl_3$]: δ 17.2 (5-CH_3), 20.7 (CH_3 (acetyl)), 20.8 (CH_3 (acetyl)), 38.9 (2'-CH_2), 63.4 (5'-CH_2), 73.8 (3'-CH), 83.3 (1'-CH), 87.6 (4'-CH), 105.8 (5-C), 145.0 (5-CH (triazolyl)), 145.5 (6-CH), 153.4 (C-4), 153.6 (3-CH (triazolyl)), 158.2 (C-2), 170.1 (CO (acetyl)), 140.3 (CO (acetyl)); Mass Spectrum (FAB$^+$): *m/z* 777 (2M+Na, 15%), 755 (2M+H, 15%), 400 (M+Na, 48%), 378 (M+H, 89%), 201 (27%), 178 (100%) [Found: *m/z* 378.1419 (M+H). $C_{16}H_{20}N_5O_6$ requires 378.1414].

1-[2'-Deoxy-5'-O-(4,4'-dimethoxytrityl)-β-D-erythro-pentofuranosyl]-pyrimidine-2,4(1H,3H)-dione 4: 2'-Deoxyuridine (930 mg 4.10 mmol) was co-evaporated with

anhydrous pyridine (3x15mL), dissolved in pyridine (40 mL) and placed under Ar to which 4,4'-dimethoxytrityl chloride (1.67 g, 4.89 mmol) was added and the mixture stirred at room temperature for 135 minutes. The reaction was quenched with MeOH and concentrated under vacuum to an oil. Flash chromatography, eluting with EtOAc, gave the title compound (2.16 g, 88%) as a solid.

1-(3',5'-Di-O-Diacetyl-2'-deoxy-β-D-erythro-pentofuranosyl)pyrimidine-2,4(1H,3H)-dione 18: 2'-Deoxyuridine (2.81 g, 12.32 mmol) was dried by co-evaporation with anhydrous pyridine (3x15 mL), dissolved in pyridine (100 mL) and placed under Ar. A catalytic amount of dimethylaminopyridine (50 mg) was added and the solution stirred at 0°C for 15 minutes. Acetic anhydride (13.6 mL, 123 mmol) was added and the solution was stirred at 4°C for 14 h before being concentrated under vacuum. Flash chromatography, eluting with EtOAc, yielded the title compound (3.11 g, 81%) as a foam; TLC (EtOAc): R_f 0.32; ^{1}H-NMR ($CDCl_3$) δ 2.11 (s, 6 H, 2xCH_3), 2.54 (m, 2 H, 2'-CH_2), 4.35 (m, 3 H, 5'-CH_2, 4'-CH), 5.22 (d, 1 H, *J* 6.5 Hz, 3'-CH), 5.79 (d, 1 H, *J* 7.3 Hz, 5-CH), 6.28 (dd, 1 H, *J* 6, 1.5 Hz, 1'-CH), 7.49 (d, 1 H, *J* 7.5 Hz, 6-CH) .

1-(3',5'-Di-O-acetyl-2'-deoxy-β-D-erythro-pentofuranosyl)-4-(1,2,4-triazolyl)-pyrimidine-2(1H)-one 22: Nucleoside **18** (3.08 g, 9.9 mmol) was dried by co-evaporated with anhydrous pyridine (3x40 mL), dissolved in pyridine (100 mL), placed under Ar and the solution cooled to 0°C. 4-Chlorophenyl phosphorodichloridate (3.2 mL, 19.7 mmol) was added and after 5 minutes 1,2,4-1*H*-triazole (2.7 g, 39.4 mmol) was added and the mixture stirred at room temperature for 72 h before being concentrated under vacuum. The residual brown oil was dissolved in CH_2Cl_2 (600 mL). The organic layer was extracted with water (3x300 mL), dried ($MgSO_4$), and the organic phase concentrated under vacuum. Flash chromatography, eluting with EtOAc, followed by crystallisation from EtOAc, gave the title compound as colourless crystals (1.89 g, 53%); m.p. 150-151°C; TLC (EtOAc): R_f 0.17; UV(95% EtOH): λ_{max} 250 nm (8200), 314 nm (4900); IR (KBr disc): υ_{max} 3121, 1751, 1675, 1670, 1552, 1515, 1239, 1059 cm^{-1}; ^{1}H NMR [$(CD_3)_2SO$]: δ 2.02 (s, 3 H, CH_3

(acetyl)), 2.07 (s, 3 H, CH_3 (acetyl)), 2.41 (m, 1 H, 2'-CH), 2.61 (m, 1 H, 2'-CH), 4.27 (d, 2 H, *J* 4.4 Hz, 5'-CH_2), 4.38 (m, 1 H, 4'-CH), 5.21 (m, 1 H, 3'-CH), 6.14 (t, 1 H, *J* 6.7 Hz, 1'-CH), 7.04 (d, 1 H, *J* 7.3 Hz, 5-CH), 8.41 (s, 1 H, 3-CH (triazolyl)), 8.45 (d, 1 H, *J* 7.3 Hz, 6-CH), 9.44 (s, 1 H, 6-CH); ^{13}C NMR [$(CD_3)_2SO$]: δ 22.1 (CH_3 (acetyl)), 22.3 (CH_3 (acetyl)), 38.9 (2-CH_2), 65.1 (5'-CH_2), 75.6 (3'-CH), 84.0 (1'-CH), 89.2 (4'-CH), 95.6 (5-CH), 145.2 (5-CH (triazolyl)), 149.5 (6-CH), 155.0 (4-C), 155.6 (6-CH (triazolyl)), 160.3 (2-C), 171.5 (CO (acetyl)), 171.6 (CO (acetyl)); Mass Spectrum (FAB^+): *m/z* 386 (M+Na, 10%), 364 (M+1, 63%), 201 (32%), 164 (100%), 81 (96%), 69 (90%); Anal. Calcd for $C_{15}H_{17}N_5O_6$: C, 49.4; H, 4.8; N, 19.2. Found: C, 49.6; H, 4.7; N, 19.3.

1-(2'-Deoxy-β-D-*erythro*-pentofuranosyl)-4N-(phenethyl)-pyrimidine-2(1H)-one ***23:***

Method A: To a solution of nucleoside **18** (1.2 g, 3.28 mmol) in dioxane (40 mL) was added phenethylamine (0.41 mL, 3.28 mmol) and the solution stirred at 80°C for 22.5 h. Evaporation of the solvent gave a foam which was dissolved in saturated methanolic ammonia (20 mL) and the solution stirred overnight before being concentrated under vacuum. Flash chromatography, eluting with EtOAc then EtOAc-MeOH 4:1 gave the title compound (1.1 g, 85%) as a foam; m.p. 44°C; TLC (EtOAc-MeOH 4:1): R_f 0.60; UV (95% EtOH): λ_{max} 274 nm (12700); IR (KBr disc): υ_{max} 3341, 3130, 3025, 2933, 2356, 2331, 1647, 1565, 1514, 1452, 1325, 1089, 1050, 943, 785, 698 cm^{-1}; ^1H NMR [$(CD_3)_2SO$]:δ 1.97 (m, 1 H, 2'-CH), 2.09 (m, 1 H, 2'-CH), 2.79 (t, 2 H, *J* 7.0 Hz, $PhCH_2$), 3.45 (d, 2 H, *J* 3 Hz, 5'-CH_2), 3.54 (s, 2 H, NCH_2), 3.76 (d, 2 H, *J* 3.1 Hz), 4.12 (s, 1 H, 3'-CH), 4.96 (t, 1 H, *J* 5.1 Hz, 5'-OH), 5.19 (s, 1 H, *J* 4 Hz, 3'-OH), 5.73 (d, 1 H, *J* 7.5 Hz, 5-CH), 6.16 (t, 1 H, *J* 6.6 Hz, 1'-CH), 7.33 (m, 5 H, 5xCH (Ph)), 7.74 (m, 2 H, NH, 6-CH); ^{13}C NMR [$(CD_3)_2SO$]: δ 34.5 (2'-CH_2), 40.2 ($PhCH_2$), 41.3 (NCH_2), 61.4 (5-CH_2), 70.5 (3'-CH), 84.9 (1'-CH), 87.2 (4'-CH), 94.2 (5-CH), 126.1, 128.4, 128.7 (3xCH, (Ph)), 139.5 (6-CH) 139.9 (C, (Ph)), 156.4 (4-C), 163.3 (2-C); Mass Spectrum (FAB^+): *m/z* (I_r) 685 (2M+Na, 5%), 663 (2M+H, 13%), 354 (M+Na, 8%), 332 (M+H, 26%), 216 (100%), [Found: *m/z* 332.1612 (M+H) $C_{12}H_{22}N_3O_4$ requires 332.1610]; Anal. Calcd. for $C_{17}H_{21}N_3O_4$. 0.5 H_2O: C, 60.0; H, 6.2; N, 12.4. Found: C, 60.4; H, 6.4; N, 12.6.

*1-(2'-Deoxy-β-D-*erythro*-pentofuranosyl)-4N-(phenethyl)-pyrimidine-2-(1H)-one 23:* Method B: To nucleoside **38** (191 mg 0.48 mmol) dissolved in dioxane (10 mL) was added phenethylamine (0.06 mL, 0.53 mmol) and the solution stirred at 80°C overnight before being concentrated under vacuum. Flash chromatography, eluting with EtOAc then EtOAc-MeOH 4:1 gave the title compound (135 mg, 85%) as a foam. Analytical data were identical to those given in Method A.

Attempted synthesis of 1-[2'-Deoxy-5'-O-(4,4'-dimethoxytrityl)-β-D-*erythro*-pentofuranosyl]-4N-phenethyl-pyrimidine-2(1H)-one 24: Method A. Nucleoside **23** (500 mg, 1.51 mmol) was dried by co-evaporation with anhydrous pyridine (3x20 mL), dissolved in pyridine (30 mL) and placed under Ar to which 4,4'-dimethoxytrityl chloride (614 mg, 1.81 mmol) was added and the reaction mixture stirred at room temperature for 5 h. The reaction was quenched with MeOH and concentrated under vacuum to an oil. Flash chromatography, eluting with EtOAc-MeOH (4:1), resulted in quantitative recovery of the starting material.

Attempted synthesis of 1-[2'-Deoxy-5'-O-(4,4'-dimethoxytrityl)-β-D-*erythro*-pentofuranosyl]-4N-phenethyl-pyrimidine-2(1H)-one 24: Method B. Nucleoside **23** (500 mg, 1.51 mmol) was dried under vacuum at room temperature overnight and dissolved in nitromethane (100 mL). 2,6-Di-*t*-butyl-4-methylpyridine (1.24 g, 0.604 mol) and 4,4'-dimethoxytrityl tetrafluoroborate (2.36 g, 604 mmol) were added and the reaction mixture refluxed for 12 h before being concentrated under vacuum. Flash chromatography, eluting with EtOAc containing Et_3N (1%), gave the title compound (28.6 mg, 3%) as an oil. ^1H NMR analysis were identical to that for Method C.

1-[2'-Deoxy-5'-O-(4,4'-dimethoxytrityl)-β-D-*erythro*-pentofuranosyl]-4N-phenethyl-pyrimidine-2(1H)-one 24: Method C. To a solution of nucleoside **42** (350 mg 0.503 mmol) in dioxane (10 mL) phenethylamine (0.063 mL, 0.503 mmol) was added and the solution

stirred at 70°C for 17 h before being concentrated under vacuum. Flash chromatography, eluting with in EtOAc containing Et_3N (1%), gave the title compound (302 mg, 95%) as a foam; m.p. 80.5°C; TLC (EtOAc): R_f 0.57; UV (95% EtOH) λ_{max} 275 nm (13000), 234 nm, (25000); IR (KBr disc): υ_{max} 3463, 3417, 3097, 3060, 2930, 2834, 2364, 2340, 1670, 1525, 1449,1388, 1290, 1249, 1180, 1117, 1029, 990 cm^{-1}; ^1H NMR [$(CD_3)_2SO$]: δ 2.06 (m, 1 H, 2'-CH), 2.16 (m, 1 H, 2'-CH), 2.80 (t, 2 H, *J* 7.3 Hz, CH_2Ph), 3.19 (s, 2 H, 5'-CH_2), 3.44 (t, 2 H, *J* 7.3 Hz, CH_2N), 3.73 (s, 6 H, 2xCH_3O), 3.86 (m,1 H, 4'-CH), 4.24 (s, 1 H, 3'-CH), 5.31 (s, 1 H, 3'-OH), 5.59 (d, 1 H, *J* 7.4 Hz, 5-CH), 6.19 (t, 1 H, *J* 6.8 Hz, 1'-CH), 6.89 (d, 4 H, *J* 8.3 Hz, 4xCH (DMTr)), 7.30 (m, 14 H, 9xCH (DMTr), 5xCH (Ph)), 7.61 (d, 1 H, *J* 7.4 Hz, 6-CH), 7.79 (t, 1 H, *J* 4.5 Hz, NH); ^{13}C NMR [$(CD_3)_2SO$]: δ 34.5 (2'-CH_2), 39.8 (CH_2Ph), 40.4 (CH_2N), 55.1 (2xCH_3O (DMTr)), 62.4 (5'-CH_2), 70.0 (3'-CH), 84.6 (1'-CH), 85.2 (4'-CH), 85.8 (Ar_3C (DMTr)), 94.6 (5-CH), 113.2 (C (Ph)), 126.2, 126.8, 127.7, 127.9, 128.4, 128.6, 129.7, 130.2 (13xCH (DMTr), 5xCH (Ph)), 135.3, 135.6 (C (DMTr)), 139.4 (4-C), 139.5 (6-CH), 144.7 (PhC (DMTr)), 158.4 (2x$\underline{C}OCH_3$ (DMTr)), 163.5 (2-C); Mass Spectrum (FAB^+): 656 (M+Na, 10%), 634 (M+H, 38%), 303 (100%), 216 (25%),105 (12%); *m/z* ; Anal. Calcd. for $C_{38}H_{39}N_3O_6$: C, 72.0; H, 6.2; N, 6.6. Found: C, 71.6; H, 6.0; N, 6.6.

1-[2'-Deoxy-3'-5'-di-O-acetyl-β-D-erythro-pentofuranosyl]-4N-propyl-[4-(3N,8N-di-t-butoxycarbonyl)spermidino]pyrimidine-2(1H)-one 25: To a solution of nucleoside **22** (300 mg, 0.862 mmol) in dioxane (4 mL) was added 4N-propyl-(*1N,8N*-di-t-butoxycarbonyl)spermidine (322 mg, 0.862 mmol) and stirred at 80°C for 20 h before being concentrated under vacuum. Flash chromatography, eluting with EtOAc-MeOH (1:1), yielded the title compound (282 mg, 49%) as a hydroscopic foam; TLC (EtOAc-MeOH 1:1): R_f 0.24; UV (95% EtOH) λ_{max} 274 nm (10300); IR (KBr disc): υ_{max} 3411, 3136, 2937, 1745, 1700, 1648, 1578, 1512, 1458, 1366, 1248, 1172, 1115, 788 cm^{-1}; ^1H NMR [$CDCl_3$]: δ 1.44 (s, 18 H, 6xCH_3), 1.49 (m, 4 H, 6-CH_2, 7-CH_2), 1.65 (m, 2 H, 2"-CH_2), 1.75 (m, 2 H, 2-CH_2), 2.08 (s, 6 H, 2xCH_3 (acetyl)), 2.55 (m, 8 H, 2'-CH_2, 1"-CH_2, 3-CH_2, 5-CH_2), 3.15 (m, 3"-CH, 8-CH), 3.58 (m, 2 H, 1-CH_2), 4.25 (m, 1 H, 4'-CH), 4.53 (m, 2 H,

5'-CH_2), 4.75 (s, 1 H, NH (Boc)), 505 (s, 1 H, NH (Boc)), 5.17 (m, 1 H, 3'-CH), 5.76 (d, 1 H, *J* 7.4 Hz, 5-CH), 6.40 (1 H, dd, *J* 5.4 Hz, 3.0 Hz, 1'-CH), 7.40 (s, 1 H, NH), 7.42 (d, 1 H, *J* 7.4 Hz, 6-CH); ^{13}C NMR [$(CD_3)_2SO$]: δ 20.7, (CH_3 (acetyl)), 20.8 (CH_3 (acetyl)), 23.9, 26.0, 26.5, 27.8 (CH_2), 28.3 (6xCH_3 (Boc)), 37.1, 38.2, 39.8 (CH_2) 40.2 (2'-CH_2), 52.1, ($\underline{C}H_2NHBoc$), 53.2 ($\underline{C}H_2NHBoc$), 63.9 (5'-CH_2), 74.3 (3'-CH), 77.9 (2x$(CH_3)_3\underline{C}$ (Boc)), 81.8 (1'-CH), 85.9 (4'-CH), 95.8 (5-CH), 137.2 (6-CH), 156.0 (2xCO (Boc)), 163.9 (2-C), 170.3 (CO (acetyl)), 170.4 (CO (acetyl)); Mass Spectrum (FAB^+): *m/z* 697 (M+H), 597 (13%), 578 (14%), 526 (12%), 497 (9%), 384 (92%), 372 (50%), 297 (14%), 226 (15%), 152 (100%), 125 (22%); Anal. Calcd. for $C_{33}H_{56}N_6O_5$ 0.5 H_2O: C, 56.2; H, 7.9; N, 11.9. Found: C, 56.2; H, 7.9; N, 11.9.

1-[2'-Deoxy-β-D-*erythro*-pentofuranosyl]-4N-propyl-[4-(3N,8N-di-t-butoxycarbonyl)-spermidino]-pyrimidine-2-(1H)-one 26: Method A. To a solution of nucleoside **38** (1 g, 2.54 mmol) in dioxane (50 mL) 4*N*-propyl-(*1N,8N*-di-t-butoxycarbonyl)spermidine (1.1 g, 2.79 mmol) was added and the solution stirred at 90°C for 36 h before being concentrated under vacuum. Flash chromatography, eluting with EtOAc-MeOH (1:1), yielded the title compound (1.1 g, 71%) as a foam; TLC (EtOAc-MeOH): R_f 0.47; UV (95% EtOH) λ_{max} 274 nm (10300); IR (KBr disc): υ_{max} 3463, 3421, 2978, 2935, 1641, 1567, 1520, 1369, 1286, 1168, 1095, 789 cm^{-1}; ^1H NMR [$(CD_3)_2SO$]: δ 1.35 (s, 18 H, 6xCH_3), 1.57-1.36 (m, 8 H, 2-CH_2, 2"-CH_2, 6-CH_2, 7-CH_2), 1.91(m, 1 H, 2'-CH), 1.97 (m, 1 H, 2'-CH), 2.32 (m, 6 H, 1"-CH_2, 3-CH_2, 5-CH_2), 2.90 (m, 4 H,3"-CH_2, 8"-CH_2), 3.10 (m, 2 H, 1-CH_2), 3.51 (m, 2 H, 5'-CH_2), 3.75 (m, 1 H, 4'-CH), 4.17 (s, 1 H, 3'-CH), 4.99 (s, 1 H, 5'-OH), 5.21 (s, 1 H, 3'-OH), 5.72 (d, 1 H, *J* 7.4 Hz, 5-CH), 6.15 (t, 1 H, *J* 6.6 Hz, 1'-CH), 6.80 (s, 2 H, 2xNH (Boc)), 7.70 (m, 2 H, 6-CH, NH); ^{13}C NMR [$(CD_3)_2SO$]: δ 23.5, 26.3, 26.9 (CH_2), 37.3 (CH_3), 37.6, 38.7 (CH_2), 39.3 (2'-CH_2), 48.8 ($\underline{C}H_2NHBoc$), 52.6 ($\underline{C}H_2NHBoc$), 61.6 (5'-CH_2), 76.1 (3'-CH), 77.4 (2x$(CH_3)_3\underline{C}$ (Boc)), 85.4 (1'-CH), 86.4 (4'-CH), 94.7 (5-CH), 139.5 (6-CH), 155.3 (4-C), 155.6 (CO, (Boc)), 163.5 (2-C); Mass Spectrum (FAB^+): *m/z* 635 (M+Na, 35%), 613 (M, 53%), 612 (M, 4%), 384 (13%), 372 (10%), 152 (100%), 125

(25%) [Found: *m/z* 613.3911 (M+H) $C_{29}H_{53}N_6O_8$ requires 613.3925]; Anal. Calcd. for $C_{21}H_{52}N_6O_8$. 0.5 H_2O: C, 56.0; H, 8.4; N, 13.5. Found: C, 56.3; H, 8.2; N, 13.1.

1-[2'-Deoxy-β-D-*erythro*-pentofuranosyl]-4N-propyl-[4-(3N,8N-di-t-butoxycarbonyl)-spermidino]-pyrimidine-2(1H)-one 26: Method B. To a solution of nucleoside **25** (193 mg, 0.227 mmol) in saturated methanolic ammonia (10 mL) was stirred overnight before being concentrated under vacuum. Flash chromatography, eluting with EtOAc-MeOH (1:1), yielded the title compound as a foam (169 mg, 99%). Analytical data were identical to those given in Method A

4,4'-Dimethoxytrityl alcohol 27: 4,4'-Dimethoxytrityl chloride (3.38 g, 10 mmol) was dissolved in THF (40 mL) to which 0.5 M aqueous NaOH (40 mL) was added and stirred at room temperature for 30 minutes. The reaction mixture was extracted with CH_2Cl_2 (80 mL) and dried with Na_2SO_4. Flash chromatography, eluting with CH_2Cl_2-MeOH (9:1), yielded a clear oil which was crystallised from light petroleum ether to give the title compound (3.18 g, 99%) as a white solid. ^{1}H NMR ($CDCl_3$): δ 2.90 (1 H, m, OH), 3.60 (6 H, s, 2xCH_3O), 6.47-7.15 (13 H, m, 13xCH (Ar)).

4,4'-Dimethoxytrityl Tetrafluoroborate 28: 4,4'-Dimethoxytrityl alcohol (1g, 3.12 mmol) was dissolved in warm acetic anhydride (6.1 mL, 64 mmol). At room temperature, 40% aqueous tetrafluoroboric acid (1.1 mL, 14.6 mL) was added slowly during 2 h at such a rate that the solution temperature did not rise above 25°C. After the solution turned dark red, dry ether was added the title product precipitated as deep orange crystals (1.16 g, 95%) which were dried under vacuum overnight. ^{1}H NMR ($CDCl_3$): δ3.95 (6 H, s, 2xOCH_3), 7.05-7.60 (13 H, m, 13xCH (Ar)).

1-[2'-Deoxy-5'-O-(4,4'-dimethoxytrityl)-3'-O-acetyl-β-D-*erythro*-pentofuranosyl]-pyrimidine-2,4(1H,3H)-dione 29: Nucleoside **4** (501 mg, 0.95 mmol) was dried by co-evaporation with anhydrous pyridine (3x15mL), dissolved in pyridine (20 mL) and placed

under Ar to which a catalytic amount of dimethylaminopyridine (5 mg) was added. Acetic anhydride (2.03 mL, 95 mmol) was added and the solution stirred overnight at 4°C before being concentrated under vacuum. Flash chromatography, eluting with EtOAc, yielded the title compound (490 mg, 90%) as a foam; m.p. 72.5°C; TLC (EtOAc): R_f 0.78; UV (95% EtOH) λ_{max} 233 nm (17300), 259 nm, (7800); IR (KBr disc): υ_{max} 3423, 2933, 1648, 1571, 1510, 1458, 1320, 1292, 1193, 1093, 1050, 942, 784, 703 cm^{-1}; ^1H NMR [$(CD_3)_2SO$]:δ 2.31 (m, 2 H, 2'-CH_2), 3.25 (m, 5 H, 5'-CH_2,CH_3 (acetyl)), 4.04 (m, 1 H, 4'-CH), 5.22 (m, 1 H, 3'-OH), 5.43 (d, 1 H, *J* 7,9 Hz, 5-CH), 6.15 (t, 1 H, *J* 6.0 Hz, 1'-CH), 6.89 (d, 4 H, *J* 8.8 Hz, 4xCH (DMTr)), 7.31 (m, 9 H, 9xCH (DMTr)), 7.61 (d, 1 H, *J* 7.9 Hz, 6-CH), 11.20 (s, 1 H, 3-NH); ^{13}C NMR [$(CD_3)_2SO$]: δ 20.9 (CH_3 (acetyl)), 38.9 (2'-CH_2), 55.2 (2xOCH_3 (DMTr)), 63.6 (5'-CH_2), 80.2 (Ar_3C (DMTr)), 83.2 (1'-CH), 84.6 (4'-CH), 86.0 (5'-CH), 113.5, 126.7, 127.0, 127.6, 128.1, 129.1 (13xCH (DMTr)), 135.3, 135.5 (2xC (DMTr)), 144.8 (PhC (DMTr)), 147.8 (6-CH), 158.1, 158.2 (2x$\underline{C}OCH_3$ (DMTr)), 163.3 (4-CO), 170.3 (2-CO), 172.3 (CO (acetyl)); Mass Spectrum (FAB^+): *m/z* 572 (M, 3%), 303 (100%), 273 (3%), 243 (4%), 215 (4%) [Found: *m/z* 572.2141 (M). $C_{32}H_{32}N_2O_8$ requires 572.2158].

1-(2'-Deoxy-β-D-*erythro*-pentofuranosyl)-4-pentafluorophenoxy-pyrimidine-2(1H)-one 38: Method A: 2'-Deoxyuridine (1.0 g, 4.39 mmol) was dried by co-evaporation with anhydrous pyridine (3x30 mL) dissolved in pyridine (40 mL), placed under Ar and the solution cooled to 0°C to. Trifluoroacetic anhydride (2.5 mL, 17.5 mmol) was added and the reaction mixture stirred at room temperature for 24 h. Pentafluorophenol (8.1 g, 43.9 mmol) in pyridine (10 mL) was added and the reaction mixture stirred for a futher 72 h at room temperature. The product mixture was diluted with water (150 mL) and the mixture extracted with EtOAc (4x200 mL). The organic phase was washed, decolourised with activated charcoal and concentrated under vacuum. The product, in EtOAc (10 mL), was precipitated into hexane and the solid collected by filtration giving the title compound (1.35 g, 78%) as a colourless powder: mp 176°C; TLC (EtOAc-MeOH 9:1): R_f 0.50; UV (95% EtOH): λ_{max} 286 nm (6700); IR (KBr disc): υ_{max} 3479, 3403, 3083, 2940, 1654, 1556, 1450, 1295, 1114 cm^{-1}; ^1H NMR [$(CD_3)_2SO$]: δ 2.07 (m, 1 H, 2'-CH), 2.29 (m, 1 H, 2'-

CH), 3.62 (s, 2 H, 5'-CH_2), 3.88 (d, 1 H, *J* 3.4 Hz, 4'-CH) 4.23 (s, 1 H, 3'-CH), 5.14 (s, 1 H, 5'-OH), 5.28 (s, 1 H, 3'-OH), 6.06 (t, 1 H, *J* 6.1 Hz, 1'-CH), 6.59 (d, 1 H, *J* 7.3 Hz, 5-CH), 8.59 (d, 1 H, *J* 7.3 Hz, 6-CH); ^{13}C NMR [$(CD_3)_2SO$]: δ 41.2 (2'-CH_2), 61.0 (5'-CH_2), 70.0 (3'-CH), 87.2 (1'-CH), 88.5 (4-CH), 93.0 (5-CH), 148.0 (6-CH), 154.0 (4-C), 168.3 (2-C); ^{19}F NMR [$(CD_3)_2SO$]: δ -14.72 (t, 2 F, *J* 20.2 Hz, 2x*Meta* C-F), -19.57 (t, 1 F, *J* 21.2 Hz, *Para* CF), -21.12 (d, 2 F, *J* 24.2 Hz, 2x*Ortho* CF); Mass Spectrum (FAB): *m/z* 417 (M+Na, 15%), 395 (M+H, 22%), 305 (14%) and 279 (100%) [Found: *m/z* 395.067 (M+H). $C_{15}H_{12}F_5N_2O_5$ requires 395.068]; Anal. Calcd for $C_{15}H_{11}F_5N_2O_5$: C, 45.7; H, 2.8; F, 24.1; N, 7.1. Found: C, 45.6; H, 2.6; F, 24.1; N, 7.1.

1-(2'-Deoxy-β-D-*erythro*-pentofuranosyl)-4-pentafluorophenoxy-pyrimidine-2(1H)-one

38: Method B: 2'-Deoxyuridine (209 mg, 0.92 mmol) was dried by co-evaporation with anhydrous pyridine (3x10 mL), dissolved in pyridine (20 mL) and placed under Ar. Trimethylsilyl chloride (0.35 mL, 2.75 mmol) was added and the mixture stirred for 30 minutes before being cooled to 0°C. 4-Chlorophenyl phosphorodichloridate (0.3 mL, 1.84 mmol) was added and the reaction mixture stirred for 24 h at room temperature to which pentafluorophenol (667 mg, 3.68 mmol) in pyridine (10 mL) was added and the reaction mixture stirred for a further 72 h before being concentrated under vacuum. Flash chromatography, eluting with EtOAc-MeOH (9:1), yielded the title compound (235 mg, 65%) as a foam. Analytical data were identical to those given in Method A.

1-(2'-Deoxy-β-D-*erythro*-pentofuranosyl)-4N-methylphenethyl-pyrimidine-2-(1H)-one

41: To a solution of nucleoside **38** (900 mg, 2.28 mmol) in dioxane (30 mL) was added *N*-phenethylmethylamine (0.4 mL, 2.74 mmol) and the solution stirred at 70°C for 14 h before being concentrated under vacuum. Flash chromatography, eluting with EtOAc-MeOH (4:1), gave the title compound (779 mg, 98%) as a foam; m.p. 164°C; TLC (EtOAc-MeOH 4:1): R_f 0.57; UV (95% EtOH) λ_{max} 281 nm (13100); IR (KBr disc): υ_{max} 3447, 3317, 3091, 3005, 2914, 1631, 1521, 1448, 1417, 1326, 1265, 1189, 1116, 1049, 783, 749, 698 cm^{-1}; ^{1}H NMR [$(CD_3)_2SO$] at 35°C: δ 1.95 (m, 1 H, 2'-CH), 2.15 (m, 1 H, 2'-CH), 2.81 (t, 2 H, *J* 7.1

Hz, CH_2Ph), 2.94 (s, 3 H, CH_3N), 3.73-3.50 (m, 4 H, CH_2N, 5'-CH_2), 3.80 (m, 1 H, 4'-CH), 4.21 (m, 1 H, 3'-CH), 4.98 (s, 1 H, 5'-OH), 5.16 (s, 1 H, 3'-OH), 6.02 (m, 1 H, (20°C, dd, *J* 7.8 Hz, 15.1 Hz), 5-CH), 6.16 (t, 1 H, *J* 6.6 Hz, (20°C, s), 1'-CH), 7.25 (m, 5 H, 5xCH (Ph)), 7.88 (s, 1 H, (20°C, dd, *J* 7.4 Hz, 20.1 Hz), 6-CH); ^{13}C NMR [$(CD_3)_2SO$] at 20°C: δ 32.6, 33.4 (CH_2Ph), 35.3, 36.3 (CH_3N), 40.2 (2'-CH_2), 50.1, 51.4 (CH_2N), 61.3 (5'-CH_2), 70.4 (3'-CH), 85.0 (1'-CH), 87.3 (4'-CH), 91.5 (5-CH), 126.4, 128.4, 128.7, 129.0 (5xCH (Ph)), 138.5, 139.2 (C (DMTr)), 140.9, 141.4 (6-CH), 154.3 (4-C), 162.7 (2-C); Mass Spectrum (FAB^+): *m/z* 368 (M+Na, 40%), 346 (M+H, 91%), 230 (100%), 105 (32%) [Found: *m/z* 346.1763 (M+H). $C_{18}H_{24}N_3O_4$ requires 346.1767]; Calcd. for $C_{18}H_{23}N_3O_4$. 0.25 H_2O: C, 61.8; H, 6.6; N, 12.0. Found: C, 61.9; H, 6.8; N, 12.1.

1-[2'-Deoxy-5'-O-(4,4'-dimethoxytrityl)-β-D-*erythro*-pentofuranosyl]-4O-pentafluorophenoxy-pyrimidine-2(1H)-one 42: Nucleoside **38** (1.00 g, 2.54 mmol) was dried by co-evaporation with anhydrous pyridine (3x40 mL), dissolved in pyridine (40 mL) and placed under Ar to which 4,4'-dimethoxytrityl chloride (1.03 g, 3.05 mmol) was added and the mixture stirred at room temperature for 5 h. The reaction was quenched with MeOH and concentrated under vacuum to an oil. Flash chromatography, eluting with 2% triethylamine in EtOAc-hexane (1:1), gave an oil which was dissolved in of EtOAc (5 mL) and precipitated into hexane (100 mL) to yield the product as a white powder (1.41 g, 80%). m.p. 104°C; TLC (EtOAc-Hexane 1:1): R_f 0.58; UV (95% EtOH) λ_{max} 283 nm (11100); IR (KBr disc): υ_{max} 3463, 3417, 3097, 3060, 2931, 2834, 2364, 2340, 1670, 1552, 1525, 1289, 1249, 1180, 1117, 991, 827 cm^{-1}; ^1H NMR [$(CD_3)_2SO$]: δ 2.18 (m, 1 H, 2'-CH), 236 (m, 1 H, 2'-CH), 3.28 (m, 2 H, 5'-CH_2), 3.73 (s, 6 H, 2xOCH_3), 3.99 (d, 1 H, *J* 3.9 Hz, 4'-CH), 4.29 (t, 1 H, *J* 4.0 Hz, 3'-CH), 5.40 (d, 1 H, *J* 4.5 Hz, 3'-OH), 6.20 (t, 1 H, *J* 6.1 Hz, 1'-CH), 6.24 (d, 1 H, *J* 7.3 Hz, 5-CH), 6.91 (d, 4 H, *J* 8.9 Hz, 4xCH (DMTr)), 7.25 (m, 9 H, 4xCH (DMTr)), 7.29 (d, 1 H, *J* 7.3 Hz, 6-CH); ^{13}C NMR [$(CD_3)_2SO$]: δ 40.5 (2'-CH_2), 55.0 (2xCH_3O (DMTr)), 62.9 (5'-CH_2), 69.6 (3'-CH), 85.9 (Ar_3C (DMTr)), 86.0 (1'-CH), 86.7 (4'-CH), 92.4 (5-CH), 112.7, 113.2, 126.4, 126.8, 127.4, 127.6, 127.7, 127.9, 128.9, 129.7, 129.8 (13xCH (DMTr)), 135.1, 135.4 (2xC (DMTr)), 140.2 (4-C), 144.6 (PhC (DMTr)),

147.2 (6-CH), 153.3 (C (PFP)), 157.8 ($\underline{C}OCH_3$ (DMTr)), 158.1 ($\underline{C}OCH_3$ (DMTr)), 168 (2-C); ^{19}F NMR [$(CD_3)_2SO$]: δ -14.78 (t, 2 F, *J* 21.1 Hz, 2x*Meta* CF), -19.63 (t, 1 F, *J* 21.2 Hz, *Para* CF), -21.52 (d, 2 F, *J* 24.2 Hz, 2x*Ortho* CF); Mass Spectrum (FAB^+): *m/z* 719 (M+Na, 86%), 696(M, 60%), 619 (8%), 570 (25%), 447 (35%), 418 (13%), 377 (90%), 303 (100%); Anal. Calcd. for $C_{36}H_{29}F_5N_2O_7$: C, 62.0; H, 4.2; F, 13.7; N, 4.0. Found: C, 62.1; H, 4.3; F, 13.7; N, 4.1.

1-[2'-Deoxy-5'-O-(4,4'-dimethoxytrityl)-β-D-*erythro*-pentofuranosyl]-4N-propyl-[4-(3N,8N-di-t-butoxycarbonyl)-spermidino]-pyrimidine-2(1H)-one 43: To a solution of nucleoside **42** (750 mg 1.07 mmol) in dioxane (10 mL) was added 4N-propyl-(*1N,8N*-di-t-butoxycarbonyl)spermidine (433 mg 1.07 mmol) was the mixture stirred at 66°C for 17 h before being concentrated under vacuum. Flash chromatography, eluting with EtOAc-MeOH (4:1) yielded the title compound (960 mg, 99%) as a glassy solid; TLC (EtOAc-MeOH 4:1): R_f 0.14; ^1H NMR [$(CD_3)_2SO$]: δ 1.58-1.90 (m, 25 H, 3xCH_2, 6xCH_3, 2'-CH), 2.19 (m, 1 H, 2'-CH), 2.50 (m, 6 H, 3xCH_2), 3.11 (m, 4 H, 2xCH_2), 3.41 (4 H, 5'-CH_2, NCH_2), 3.78 (s, 6 H, 2xOCH_3), 4.18 (m, 1 H, 4'-CH), 4.50 (m, 1 H, 3'-CH), 4.91 (s, 1 H, NHBoc), 5.24 (s, 1 H, NHBoc), 5.57 (d, 1 H, *J* 7.5 Hz, 5-CH), 6.38 (t, 1 H, *J* 4.4 Hz, 1'-CH), 6.81 (d, 4 H, *J* 8.3, CH (DMTr)), 7.31 (m, 10 H, 9xCH (DMTr), NH), 7.65 (d, 1 H, *J* 7.5, 6-CH); ^{13}C NMR [$(CD_3)_2SO$]: δ 22.2, 23.1, 27.7, 28.3 (CH_2), 29.5 (6xCH_3), 39.4, 39.7, 41.2 (CH_2), 46.8 (2'-CH_2), 51.8, 54.5 ($NHCH_2$), 57.1 (2xOCH_3 (DMTr)), 65.1 (5-CH_2), 71.3 (3'-CH), 78.5 (Ar_3C (DMTr)), 85.3 (1'-CH), 86.2 (4'-CH), 87.0 (2x(CH_3)$\underline{C}$), 96.2 (5-CH), 114., 129.5, 130.2, 130.6, 131.2 (13xCH (DMTr)), 136.6, 136.8 (2xC (DMTr)), 147.1 (PhC (DMTr)), 156.1 (2xCO (Boc)), 159.9 (2-C); Mass Spectrum (FAB^+): *m/z* 937 (M+K, 70%), 591 (96%), 384 (81%), 303 (100%).

1-[2'-Deoxy-5'-O-(4,4'-dimethoxytrityl)-β-D-*erythro*-pentofuranosyl]-4N-methylphenethyl-pyrimidine-2(1H)-one 44: To a solution of nucloeoside **42** (1.0 g, 2.54 mmol) in dioxane (50 mL) was added *N*-methylphenethylamine (0.41 mL, 2.79 mmol) and the solution stirred at 70°C for 14 h before being concentrated under vacuum. Flash

chromatography, eluting with EtOAc-MeOH (9:1), yielded the title compound (600 mg, 37%) as a foam; m.p. 86°C; TLC (EtOAc-MeOH 9:1): R_f 0.53; UV (95% EtOH) λ_{max} 282 nm (17500); IR (KBr disc): υ_{max} 3433, 2935, 2362, 2339, 1638, 1523,1505, 1448,1407, 1301, 1253, 1174, 1099, 1035, 831, 703 cm^{-1}; ^1H NMR [$(CD_3)_2SO$]:δ 2.07 (m, 1 H, 2'-CH), 2.17 (m, 1 H, 2'-CH), 2.77 (m, 2 H, CH_2Ph), 2.90 (s, 1.5 H, CH_3N), 2.96 (s, 1.5 H, CHN_3), 3.21 (m, 2 H, CH_2N), 3.56 (m, 1 H, 5'-CH), 3.72 (m, 7 H, 2xCH_3O, 5'-CH), 3.87 (s, 1 H, 4'-CH), 4.25 (m, 1 H, 3'-CH), 5.31 (d, *J* 7.6 Hz, 5-CH), 6.12 (m, 1 H, 1'-CH), 6.90 (d, 4 H, *J* 8.8 Hz, 4xCH (DMTr)), 7.34 (m, 14 H, 9xCH (DMTr), 5xCH (Ph)), 7.66 (m, 1 H, 6-CH); ^{13}C NMR [$(CD_3)_2SO$]: δ 32.9 (CH_2Ph), 35.8 (CH_3N), 40.5 (2'-CH_2), 55.0 (2xOCH_3 (DMTr)), 63.4 (5'-CH_2), 70.4 (3'-CH), 85.0 (1'-CH), 85.4 (Ar_3C (DMTr)), 85.7 (4'-CH), 91.8 (5-CH), 113.2, 126.2, 127.7, 128.4, 129.8 (13xCH (DMTr), 5xCH (Ph)), 135.1, 135.4 (C (DMTr)), 138.8 (C (Ph)), 141.0 (6-CH), 144.8 (PhC (DMTr)), 158.1 (2x$\underline{C}OCH_3$ (DMTr)), 162.4 (2-C); Mass Spectrum (FAB^+): *m/z* 670 (M+Na, 15%), 648 (M+H, 14%), 344 (12%), 303 (100%), 230 (65%), 165 (10%) [Found: *m/z* 648.3091. (M+H) $C_{39}H_{42}N_3O_6$ requires 648.3074]; Anal. Calcd. for $C_{39}H_{42}N_3O_6$: C, 72.3; H, 6.3; N, 6.5. Found: C, 71.8; H, 6.4; N, 6.6.

1-[2'-Deoxy-5'-O-(4,4'-dimethoxytrityl)-β-D-*erythro*-pentofuranosyl]-4-O-pentafluorophenoxy-pyrimidine-2(1H)-one-3'-[(2-cyanoethyl)-N,N-diisopropyl]-phosphoramidite 45: Nucleoside **42** (500 mg 0.718 mmol) was dried by co-evaporation with anhydrous THF (3x10 mL), dissolved in THF (10 mL) and placed under Ar. Anhydrous triethylamine (0.29 mL, 2.16 mmol) was added followed by 2-cyanoethyl-*N,N*-diisopropyl chlorophosphoramidite (0.25 mL, 1.44 mmol) and the mixture stirred at room temperature for 3 h before being concentrated under vacuum at 25°C. Flash chromatography, eluting with in EtOAc-hexane (1:1) containing Et_3N gave the title compound (645 mg, 100%) as a foam; TLC (2% Et_3N in EtOAc-Hexane, 1:1): R_f 0.60; IR (KBr disc): υ_{max} 3436, 2970, 1679, 1629, 1517, 1453, 1284, 1249, 1176, 1026, 991, 831, 781, 725, 706 cm^{-1}; ^1H NMR [$(CD_3)_2SO$]: δ 1.15 (m, 12 H, 4xCH3 (^{i}Pr)), 1.27 (m, 1 H, 2'-CH), 2.31 (m, 1 H, 2'-CH), 2.45 (t, 1 H, *J* 6.4 Hz, CHO), 2.63 (t, 1 H, *J* 6.4 Hz, CHO), 2.75

(m, 1 H, 2'-CH), 3.49-3.73 (m, 6 H, 5'-CH_2, $CNCH_2$, 2xCH (^iPr)), 4.14 (s, 1 H, 4'-CH), 4.66 (s, 1 H, 3'-CH), 5.83 (m, 1 H, 1'-CH), 6.23 (m, 1 H, 5-CH), 6.85 (m, 4 H, 4xCH (DMTr)), 7.18-7.38 (m, 9 H, 9xCH (DMTr)), 8.41 (d, 0.5 H, 7.2 Hz, 6-CH), 8.46 (d, 0.5 H, 7.2 Hz 6-CH); ^{13}C NMR [$(CD_3)_2SO$]: δ 16.2 (CH_2N), 24.1, 24.2, 24.3, 24.4 (4xCH_3 (Pr^i)), 42.5, 42.7 (2xCH (Pr^i)), 55.0, 55.2 (2xOCH_3 (DMTr)), 58.2 (CH_2O), 67.0 (5'-CH_2), 72.0 (3'-CH), 85.9 (Ar_3C (DMTr)), 86.0 (1'-CH), 87.1 (4'-CH), 92.5 (5-CH), 113.2 (CH (DMTr)), 118.9 (CN), 125.0, 126.8, 127.6, 127.8, 129.8 (9xCH (DMTr)), 135.0 (2xC (DMTr)), 144.5 (CPh (DMTr)), 147.5 (6-CH), 153.2 (4-C), 162 (2x$\underline{C}OCH_3$ (DMTr)), 168.4 (2-C); ^{19}F NMR [$(CD_3)_2SO$]: δ -14.87 (t, 2 F, *J* 21.3 Hz, 2x*Meta* CF), -19.75 (t, 1 F, *J* 23.1 Hz, *Para* CF), -24.15 (d, 2 F, *J* 20.7 Hz, 2x*Ortho* CF); ^{31}P NMR [$(CD_3)_2SO$]: δ 43.23 ,43.74 ; Mass Spectrum (FAB^+): *m/z* 935 (M+K, 7%), 919 (M+Na, 62%), 897 (M+H, 7%), 679 (93%), 619 (35%), 593 (8%), 577 (22%), 375 (7%), 303 (100%) [Found: *m/z* 919.2833 (M+Na). $C_{45}H_{46}N_4O_8F_5NaP$ requires 919.2871]; Anal. Calcd. for $C_{45}H_{46}F_5N_4O_8P$: C, 60.3; H, 5.3; F, 10.6; N, 6.3; P, 3.5; Found: C, 60.7; H, 5.3; F, 10.1; N, 6.2; P, 3.5.

1-[2'-Deoxy-5'-O-(4,4'-dimethoxytrityl)-β-D-*erythro*-pentofuranosyl]-4N-phenethyl-pyrimidine-2(1H)-one-3'-[(2-cyanoethyl)-N,N-diisopropyl]-phosphoramidite 46: Nucloeoside **24** (460 mg, 0.79 mmol) was dried by co-evaporation with anhydrous THF (3x10 mL), dissolved in THF (10 mL) and placed under Ar. Anhydrous triethylamine (0.28 mL, 1.98 mmol) was added followed by 2-cyanoethyl-*N,N*-diisopropyl chlorophosphoramidite (0.15 mL, 0.87 mmol) and the mixture stirred at room temperature for 2.5 h before being concentrated under vacuum at 25°C. Flash chromatography, eluting with EtOAc-MeOH (9:1) containing Et_3N (2%), gave the title compound (535 mg, 87%) as a foam; TLC (2% Et_3N in EtOAc-MeOH, 9:1): R_f 0.56; IR (KBr disc): υ_{max} 3452, 2966, 2931, 2389, 2250, 1648, 1503, 1466, 1410, 1365, 1305, 1250, 1178, 1031, 982, 827, 700 cm^{-1}; ^1H NMR [$(CD_3)_2SO$]:δ 1.20 (m, 12 H, 4xCH_3 (^iPr)), 2.32 (m, 2 H, 2'-CH_2), 2.73 (t, 1 H, *J* 5.9 Hz, CNCH), 2.83 (m, 3 H, CNCH, $PhCH_2$), 3.27 (m, 2 H, 5'-CH_2), 3.44-3.68 (m, 6 H, CH_2N, CH_2O, 2xCH (^iPr)), 3.73 (s, 6 H, 2xCH_3O), 4.01 (m, 1 H, 4'-CH), 4.5 (m, 1 H, 3'-CH), 5.59 (d, 1 H, *J* 7.5 Hz, 5-CH), 6.17 (m, 1 H, 1'-CH), 6.69 (m, 4 H, 4xCH (DMTr)),

7.30 (m, 14 H, 9xCH (DMTr), 5xCH (Ph)), 7.61 (m, 1 H, 6-CH), 7.81 (t, 1 H, *J* 4.5 Hz, NH); ^{13}C NMR [$(CD_3)_2SO$]: δ 19.8 ($\underline{C}H_2CN$), 22.60 (CH_3 (^{i}Pr)), 24.10 (CH_3 (^{i}Pr)), 24.22 (CH_3 (^{i}Pr)), 24.32 (CH_3 (^{i}Pr)), 41.2 (2'-CH_2), 42.5 (CH (^{i}Pr)), 42.7 (CH (^{i}Pr)), 45.6 (CH_2N), 53.6 (2xOCH_3 (DMTr)), 58.2 (CH_2O), 62.9 (5'-CH_2), 72.8 (3'-CH), 83.5 (1'-CH), 84.8 (4'-CH), 85.8 (Ar_3C (DMTr)), 94.8 (5-CH), 118.9 (CN), 126.1, 126.8, 127.7, 128.3, 128.6, 129.7 (13xCH (DMTr), 5xCH (Ph)), 135.2, 135.3 (2xC (DMTr)), 139.4 (C (Ph)), 139.7 (6-CH), 144.6 (PhC (DMTr)), 154.9 (4-C), 158.1 (2x$\underline{C}OCH_3$ (DMTr)), 163.3 (2-C); ^{31}P NMR [$(CD_3)_2SO$]:δ 147.64, 148.08; Mass Spectrum (FAB$^+$): *m/z* 856 (M+Na, 25%), 834 (M ,11%), 616 (18%), 303 (100%), 216 (37%); Anal. Calcd. for $C_{47}H_{57}N_5O_7P$ 1.25 H_2O: C, 65.8; H, 6.7; N, 8.2. Found: C, 65.4; H, 6.9; N, 8.6.

1-[2'-Deoxy-5'-O-(4,4'-dimethoxytrityl)-β-D-*erythro*-pentofuranosyl]-4N-phenethyl-N-methyl-pyrimidine-2-(1H)-one-3'-[(2-cyanoethyl)-N,N-diisopropyl]-phosphoramidite 48: Nucloeoside **44** (916 mg, 1.42 mmol) was dried by co-evaporation with anhydrous THF (3x10 mL), dissolved in THF (10 mL) and placed under Ar. Anhydrous diisopropylethylamine (0.74 mL, 4.25 mmol) was added followed by 2-Cyanoethyl-*N,N*-diisopropyl chlorophosphoramidite (0.47 mL, 2.12 mmol) and the mixture stirred at room temperature for 2.5 h before being concentrated under vacuum at 25°C. Flash chromatography, eluting with EtOAc containing Et_3N, gave the title compound (535 mg, 87%) as a foam; TLC (2% Et_3N in EtOAc): R_f 0.32; IR (KBr disc): υ_{max} 3452, 2966, 2931, 2389, 2250, 1648, 1504, 1465, 1410, 1364, 1305, 1250, 1178, 1031, 982, 827, 700 cm^{-1}; ^{1}H NMR [$(CD_3)_2SO$]: δ 1.11 (m, 12 H, 4xCH_3 (^{i}Pr)), 2.06 (m, 1 H, 2'-CH), 2.49 (m, 1 H, 2'-CH), 2.67 (t, 1 H, *J* 5.9 Hz, CNCH), 2.84 (m, 3 H, CNCH, CH_2Ph), 2.90 (s, 1.5 H, CH_3), 2.96 (s, 1.5 H, CH_3N), 3.30 (m, 2 H, 5'-CH), 3.44-3.71 (m, 6 H, CH_2N, CH_2O, 2xCH (^{i}Pr)), 3.79 (s, 6 H, 2xCH_3O), 4.01 (dd, 1 H, *J* 6.6, 7.2 H, 4'-CH), 4.49 (m, 1 H, 3'-CH), 5.78 (d, 1 H, *J* 7.1 Hz, 5-CH), 6.15 (t, 1 H, *J* 6.5 Hz, 1'-CH), 6.90 (m, 4 H, 4xCH (DMTr)), 7.30 (m, 14 H, 9xCH (DMTr), 5xCH (Ph)), 7.61 (m, 1 H, 6-CH); ^{13}C NMR [$(CD_3)_2SO$]: δ 19.8 (CH_2), 20.9 (CH_3), 24.3 (CH_3), 32.1, 33.4 (CH_2Ph), 35.2, 36.1 (CH_3N), 39.5 (2'-CH_2), 42.5, 42.7 (2xCH (^{i}Pr)), 50.0, 51.2 (CH_2N), 55.1 (2xOCH_3 (DMTr)),58.1 (CH_2O), 82.8 5'-

CH), 72.4 (3'-CH), 83.8 (1'-CH), 85.0 (4'-CH), 91.7 (5'-CH), 113.2 (4xCH (DMTr)), 118.4 (CN), 127.9, 128.4, 128.9, 129.5, 129.8 (9xCH (DMTr),5xCH (Ph)), 135.3 (2xC (DMTr)), 138.2 (Ć (Ph)), 139.0 (6-CH), 144.2 (PhC (DMTr)), 154.0 (4-C), 158.2 (2x$\underline{C}OCH_3$ (DMTr)), 162.1 (2-C); ^{31}P NMR [$(CD_3)_2SO$]:δ 147.67, 148.10; Mass Spectrum (FAB^+): *m/z* 870 (M+Na, 15%), 848 (M, 9%), 630 (24%), 303 (100%), 230 (52%) [Found: *m/z* 848.4157 (M). $C_{48}H_{59}N_5O_7P$ requires 848.4152].

1-Chloro-2'-Deoxy-3,5-di-O-toluoyl-α-D-erythro-pentofuranose 61: Acetic acid was cooled to 0°C, saturated with HCl gas and sugar **66** (10 g, 0.026 mmol) was added. Over a period of 10 minutes the title compound (9.8 g, 91%) precipitated as a yellow powder and was collected by filtration, washed with diethylether and dried under vacuum with P_2O_5; TLC (EtOAc-Hexane 1:2): R_f 0.44; ^1H NMR [$(CD_3)_2SO$]:δ 2.40 (s, 3 H, CH_3), 2.41 (s, 3 H, CH_3), 2.83 (m, 2 H, 2'-CH_2), 4.65 (qd, 2 H, *J* 12.1 Hz, 4.2 Hz, 5'-CH_2), 4.85 (m, 1 H, 4'-CH), 5.54 (dd, 1 H, *J* 6.6 Hz, 2.7 Hz, 3'-CH), 6.46 (d, 1 H, *J* 4.8 Hz, 1'-CH), 7.24 (t, 4 H, *J* 7.6, 4-CH (Ar)), 7.88 (d, 4 H, *J* 8.1 Hz, 4xCH (Ar)), 7.98 (d, 4 H, *J* 8.1, 4xCH (Ar)).

2-Chloro-5-phenylmethoxy-pyridine 63: 6-Chloro-2-pyridinol (1.0 g, 7.72 mmol) and silver carbonate (3.2 g, 11.6 mmol) were dispersed in toluene (50 mL) and heated to 110°C for 1 h. The reaction mixture was cooled to 70°C to which benzyl bromide (1.85 mL, 15.4 mmol) was added before the temperature was returned to 110 °C and the reaction mixture stirred overnight. After cooling, the solid residue was filtered through celite and the filtrate concentrated under vacuum. Flash chromatography, eluting with hexane, yielded the title compound (1.07 g, 63%) as a colourless liquid; TLC (Hexane): R_f 0.59; UV (95% EtOH) λ_{max} 278 nm (5600); IR (liquid film): υ_{max} 3091, 3862, 3029, 2944, 2885, 2561, 2365, 1951, 1749, 1589, 1564, 1494, 1436, 1371, 1296, 1265, 1159, 1069, 989, 917, 871, 786, 748 cm^{-1}; ^1H NMR [$(CD_3)_2SO$]: δ 5.31 (s, 2 H, CH_2), 6.85 (d, 1 H, *J* 7.7 Hz, 5-CH), 7.07 (d, 1 H, *J* 7.7 Hz, 3-CH), 7.37 (m, 5 H, Ar 5xCH), 7.71 (d, 1 H, *J* 7.6 Hz, 4-CH); ^{13}C NMR [$(CD_3)_2SO$]: δ 69.0 (CH_2), 109.5 (5-CH), 116.7 (3-CH), 128.1 (Ar 5xCH), 136.4 (C-Cl), 141.8 (4-CH), 147.1 (Ph C), 162.7 (6-C); Mass Spectrum (FAB^+): *m/z* 219 (M,31%), 91

(100%), 65 (25%), 51 (7%), 39 (15%) [Found: *m/z* 219.0455 (M). $C_{12}H_{10}ClNO$ requires 219.0451].

2'-Deoxy-1'-O-methyl-α/β-D-erythro-pentofuranose 65: Acetylchloride (1 mL, 14.6 mmol) was slowly added to MeOH (50 mL). The resulting acidic solution was slowly added to a solution of 2'-deoxyribose (30 g, 0.22 mol) in MeOH (450 mL). After 10 minutes Ag_2CO_3 (9.8 g, 35.5 mmol) was added and the reaction mixture was filtered through celite and concentrated under vacuum to yield a the crude product as a colourless oil in (21.2 g, 65%) yield; TLC (EtOAc-MeOH 4:1): R_f 0.54; ^{1}H NMR [$(CD_3)_2SO$]:δ 1.62 (m, 1 H, 2'-CH), 1.91 (m, 1 H, 2'-CH), 3.27 (s, 3 H, OCH_3), 3.40 (m, 1 H, 4'-CH), 3.66 (m, 2 H, 5-CH_2), 3.90 (m, 1 H, 3'-CH), 4.64 (s, 1 H, 5'-OH), 4.88 (s, 1 H, 3'-OH), 4.93 (m, 2 H, 1'-CH, 3'-OH).

2'-Deoxy-1'-3,5-di-O-toluoyl-O-methyl-α/β-D-erythro-pentofuranose 66: Sugar **65** (10 g, 0.06 mmol) was dissolved in pyridine (80 mL), cooled to 0°C, to which p-toluoyl chloride (17.6 mL, 13.4 mmol) was added in four portions. The reaction was stirred for 24 h before being concentrated under vacuum. The oil was dissolved in water (500 mL) and partitioned into diethylether (500 mL) three times. The organic layer was concentrated under vacuum (250 mL) then washed with 0.1 M HCl aqueous acid (250 mL) three times followed by washing with saturated sodium hydrogen carbonate aqueous solution (250 mL) three times. The organic layer was concentrated under vacuum to give the crude product **66** (6.0 g, 68%) as a yellow oil; TLC (EtOAc-hexane 1:2): R_f 0.44; ^{1}H NMR [$(CD_3)_2SO$]:δ 2.30 (m, 7 H, 2'-CH, 2xCH_3), 2.56 (m, 1 H, 2'-CH), 3.33 (s, 3 H, OCH_3), 4.56 (m, 3 H, 4'-CH, 5'-CH_2), 5.19 (m, 1 H, 3'-CH), 5.4 (m, 1 H, 1'-CH), 7.25 (m, 4 H, 4xH (Ar)), 7.92 (m, 4 H, 4xH (Ar)).

Attempted synthesis of 2-chloro-5-(2'-Deoxy-β-D-erythro-pentofuranosyl)-6-phenmethoxy-pyridine 67 and 2-chloro-3-(2'-Deoxy-β-D-erythro-pentofuranosyl)-6-phenmethoxy-pyridine 68: Butyl lithium (1.3 M in hexane, 12.3 mL, 32.0 mmol) was

added to a solution of isopropylamine (1.30 mL, 35.0 mmol) at 0°C under Ar in THF (100 mL) and the mixture stirred. After 30 minutes the reacton mixture was cooled to -78°C and a solution of 2-chloro-5-phenmethoxy-pyridine (3.84 g, 32.0 mmol) in THF (50 mL) was added. The mixture was stirred for 1.5 h before sugar (13.3 g, 32.0 mmol) was added and the reaction mixture stirred. After 1 h the reaction temperature was allowed to rise to -40°C over 90 minutes before the reaction was quenched with MeOH. The reaction mixture was concentrated under vacuum then dissolved in saturated methanolic ammonia (70 mL). After stirring at room temperature for 24 h, TLC analysis showed two UV active spots and staining with ethanolic anisaldehyde solution indicated that these products contained a sugar. The products were isolated by flash chromatography but ^{1}H NMR showed that they were not the title compounds; their identity remains unknown.

1-(5'-O-Acetyl-3'-azido-2',3'-dideoxy-β-D-*erythro*-pentofuranosyl)-5-methyl-pyrimidine-2,4(1H,3H)-dione 89: AZT (1 g, 3.74 mmol) was dried by co-evaporation with anhydrous pyridine (3x15 mL), dissolve in pyridine (40 mL) and placed under Ar. A catalytic amount of dimethylaminopyridine (30 mg) was added and the solution stirred at 0°C for 15 minutes. Acetic anhydride (4.14 mL, 37.4 mmol) was added and the solution stirred at 4°C for 14 h before being concentrated under vacuum. Flash chromatography, eluting with EtOAc, yielded the title compound as a foam (1.01 g, 81%); TLC (EtOAc): R_f 0.51; ^{1}H NMR [$(CD_3)_2SO$]:δ 1.78 (s, 3 H, 5-CH_3), 2.06 (s, 3 H, acetyl CH_3), 2.28 (m, 1 H, 2'-CH), 2.45 (m, 1 H, 2'-CH), 3.95 (m, 1 H, 3'-CH), 4.23 (m, 2 H, 5'-CH_2), 4.95 (m, 1 H, 4'-CH), 6.12 (t, 1 H, J 6.2 Hz, 1'-CH), 4.93 (s, 1 H, 6-CH), 11.38 (s, 1H, NH).

1-(2'-Deoxy-β-D-*erythro*-erythro-pentofuranosyl)-4-(1,2,4-triazolyl)-pyrimidine-2(1H)-one 90: 2'-Deoxyuridine (1.00 g, 4.40 mmol) was dried by co-evaporated with anhydrous pyridine (3x30 mL), dissolved in pyridine (40 mL), placed under Ar and the solution cooled to 0°C. Trifluoroacetic anhydride (1.86 mL, 13.16 mmol) was slowly added and the solution stirred at room temperature for 24 h to which 1,2,4-1*H*-triazole (3.03 g, 43.86 mmol) was added and the mixture stirred at room temperature for 72 h before being

concentrated under vacuum. Flash chromatography, eluting with CH_2Cl_2-CH_3OH (4:1), yielded the title compound (904 mg, 79%) as a colourless solid; m.p. >300°C; TLC (CH_2Cl_2-CH_3OH 4:1): R_f 0.76; UV (95% EtOH): λ_{max} 249 nm (8200), 315 nm (4900); IR (KBr disc): υ_{max} 3303, 3145, 3106, 2951, 2900, 1660, 1550, 1515, 1451, 1425, 1282, 1184, 1060 cm^{-1}; ^1H NMR [$(CD_3)_2SO$]: δ 2.10 (m, 1 H, 2'-CH), 2.38 (m, 1 H, 2-CH), 3.62 (m, 2 H, 5'-CH_2), 3.91 (m, 1 H, 4'-CH), 4.23 (m, 1 H, 3'-CH), 5.13 (t, 1 H *J* 5.1 Hz, 5'-OH), 5.31 (d, 1 H, *J* 4.3 Hz, 3'-OH), 6.10 (t, 1 H, *J* 6.0 Hz, 1'-CH), 6.96 (d, 1 H, *J* 7.2 Hz, 5-CH), 8.38 (s, 1 H, 3-CH (triazolyl)), 8.69 (d, 1 H, *J* 7.2 Hz, 6-CH), 9.41 (s, 1 H, 5-CH (triazolyl)); ^{13}C NMR [$(CD_3)_2SO$]: δ 41.2 (2'-CH_2), 60.9 (5'-CH_2), 69.8 (3'-CH), 87.4 (1'-CH), 88.5 (4'-CH), 93.8 (5-CH), 143.8 (5-CH (triazolyl)), 148.3 (6-CH), 153.8 (4-C), 154.3 (6-CH (triazolyl)), 158.7 (2-C); Mass Spectrum (FAB$^+$): *m/z* 302 (M+Na, 28%), 280 (M+H, 66%), 260 (24%), 164 (67%), 150 (32%), 121 (83%), 108 (79%), 105 (65%); Anal. Calcd for $C_{11}H_{13}N_5O_4$: C, 47.3; H, 4.7; N, 25.1. Found: C, 47.2; H, 4.6; N, 24.7.

1-(3',5'-Di-O-acetyl-2'-deoxy-β-D-*erythro*-pentofuranosyl)-4-pentafluorophenoxy-pyrimidine-2-(1H)-one 91: Nucleoside **18** (1.00 g, 3.18 mmol) was dried by co-evaporation with anhydrous pyridine (3x30 mL), dissolved in pyridine (40 mL), placed under Ar and the solution cooled to 0°C. 4-Chlorophenyl phosphorodichloridate (0.99 mL, 6.37 mmol) was added and the reaction mixture stirred at room temperature for 24 h. Pentafluorophenol (2.34 g, 12.73 mmol) in pyridine (10 mL) was added and the reaction mixture stirred for 72 h before being concentrated. Flash chromatography, eluting with EtOAc-Hexane (4:1), yielded the title compound as a foam (0.93 g, 61%); m.p. 136°C; TLC (EtOAc-Hexane, 4:1): R_f .51; UV (95% EtOH) λ_{max} 297 nm (8000); IR (KBr disc): υ_{max} 3465, 3124, 2987, 2933, 2667, 2451, 1745, 1661, 1550, 1452, 1378, 1297, 1199, 1112, 1072, 1023, 989, 775 cm^{-1}; ^1H NMR [$(CD_3)_2SO$]: δ 2.06 (s, 6 H, 2xCH_3 (acetyl)), 2.41 (m, 2 H, 2'-CH_2), 4.25 (m, 2 H, 5-CH_2), 4.31 (m, 1 H, 4'-CH), 5.17 (m, 1 H, 3'-CH), 6.07 (t, 1 H ,*J* 6.5 Hz, 1'-CH), 6.62 (d, 1 H, *J* 7.3 Hz, 5-CH), 8.32 (d, 1 H, *J* 7.3 Hz, 6-CH); ^{13}C NMR [$(CD_3)_2SO$]: δ 20.6 (CH_3 (acetyl)), 20.8 (CH_3 (acetyl)), 37.3 (2'-CH_2), 63.6 (5'-CH_2), 74.1 (3'-CH), 82.4 (1'-CH), 87.4 (4'-CH), 93.2 (5-CH), 147.7 (6-CH), 153.5 (4-C), 168.7 (2-CO) 170.1 (CO

(acetyl)), 170.2 (CO (acetyl)); ^{19}F NMR [$(CD_3)_2SO$] δ;-20.8 (t, 2 F, *J* 20.2 Hz, 2x*Meta* CF), -14.8 (t, 1 F, *J* 21.2 Hz, *Para* CF), -10.4 (d, 2 F, *J* 24.2 Hz, 2x*Ortho* CF); Mass Spectrum (FAB$^+$): *m/z* 979 (2M+Na, 17%), 957 (2M+H, 27%), 501 (M+Na, 66%), 479 (M+H, 54%), 279 (100%), 201 (75%) [Found: *m/z* 479.0875 (M+H). $C_{19}H_{16}N_2O_7F_5$ requires 479.0877]. Anal. Calcd for $C_{19}H_{16}F_5N_2O_5$: C, 47.7; H, 3.1; N, 5.9. Found: C, 48.0; H, 3.2; N, 5.8.

Attempted synthesis of 1-(2'-deoxy-β-D-erythro-pentofuranosyl)-5-methyl-4-pentafluorophenoxy-pyrimidine-2-(1H)-one 92: Method A. 2'-Deoxythymidine (1.00 g, 4.13 mmol) was dried by co-evaporation with pyridine (3x30 mL), dissolved in pyridine (50 mL), placed under Ar and the solution cooled to 0°C. Trifluoroacetic anhydride (2.33 mL, 16.5 mmol) was added and the reaction mixture stirred at room temperature for 24 h. Pentafluorophenol (7.60 g, 41.3 mmol) in pyridine (10 mL) was added and the reaction mixture stirred for a further 72 h before being concentrated under vacuum to give a dark brown syrup. TLC indicated mainly starting material was present.

Attempted synthesis of 1-(2'-deoxy-β-D-erythro-pentofuranosyl)-5-methyl-4-pentafluorophenoxy-pyrimidine-2(1H)-one 92: Method B. 2'-Deoxythymidine (1.00 g, 4.13 mmol) was dried by co-evaporation in pyridine (3x30 mL) dissolved in pyridine (50 mL) and placed under Ar. Trimethylsilyl chloride (1.15 mL, 9.08 mmol) was added and the reaction mixture stirred for 4 h before being cooled to 0°C. 4-Chlorophenyl phosphorodichloridate (2.96 mL, 18.2 mmol) was added and the reaction mixture stirred for 24 h at room temperature. Pentafluorophenol (6.67 g, 36.3 mmol) in pyridine (10 mL) was added and the mixture stirred for a further 72 h. Water (10 mL) was added and the reaction mixture stirred for 1 minute before being concentrated under vacuum. TLC indicated a number UV active spots including starting material. Any attempt to isolate the title compound by flash chromatography failed.

Attempted synthesis of 1-(2'-deoxy-β-D-erythro-pentofuranosyl)-5-methyl-4-(1,2,4-triazolyl)-pyrimidine-2(1H)-one 93: Method A. 2'-Deoxythymidine (1.00 g, 4.13 mmol) was dried by co-evaporation in pyridine (3x30 mL), dissolved in pyridine (50 mL), placed under Ar and the solution cooled to 0°C. Trifluoroacetic anhydride (2.33 mL, 16.5 mmol) was added and the reaction mixture stirred at room temperature for 24 h. 1,2,4-1*H*-Triazole (2.28 g, 33.0 mmol) was added and the reaction mixture stirred at room temperature for a further 72 h before being concentrated under vacuum to give a dark brown syrup. TLC indicated mainly starting material was present and did not show the characteristic triazolyl-nucleoside spot expected under UV.

Attempted synthesis of 1-(2'-deoxy-β-D-erythro-pentofuranosyl)-5-methyl-4-(1,2,4-triazolyl)-pyrimidine-2(1H)-one 93: Method B. 2'-Deoxythymidine (1.00 g, 4.13 mmol) was dried by co-evaporation in pyridine (3x30 mL), dissolved in pyridine (50 mL) and placed under Ar to which trimethylsilyl chloride (1.15 mL, 9.08 mmol) was added and the mixture stirred for 4 h before being cooled to 0°C. 4-Chlorophenyl phosphorodichloridate (2.96 mL, 18.2 mmol) was added and the mixture stirred at room temperature for 24 h to which 1,2,4-1*H*-triazole (2.05 g, 36.3 mmol) was added and the reaction mixture stirred for a further 72 h. Water (10 mL) was added and stirred for 1 minute before being concentrated under vacuum to give a brown syrup. TLC indicated a number of UV active spots including starting material. The fluorescent spot characteristic of a triazyl-nucleoside was seen under UV but any attempts to isolated the title compound by flash chromatography failed.

1-(3',5'-Di-O-acetyl-2'-deoxy-β-D-erythro-pentofuranosyl)-5-methyl-4-pentafluorophenoxy-pyrimidine-2(1H)-one 94: Nucleoside **1** (1.00 g, 3.09 mmol) was dried by co-evaporation with anhydrous pyridine (3x30 mL), dissolved in pyridine (40 mL), placed under Ar and the solution cooled to 0°C. 4-Chlorophenyl phosphorodichloridate (1.00 mL, 6.17 mmol) was added and the reaction mixture stirred at room temperature for 24 h Pentafluorophenol (2.27 g, 12.3 mmol) was added and the reaction mixture stirred for a further 72 h before being concentrated under vacuum. Flash chromatography, eluting with

EtOAc-Hexane (1:1), yielded the title compound (0.99 g, 67%) as a foam; m.p. 126°C; TLC (EtOAc-Hexane, 1:1): R_f 0.55; UV (95% EtOH) λ_{max} 297 nm (7980); IR (KBr disc): υ_{max} 3548, 3471, 3413, 3123, 2987, 2932, 2663, 2457, 2358, 1739, 1666, 1550, 1515, 1452, 1378, 1324, 1230, 1223, 1029, 989, 775 cm^{-1}; ^1H NMR [($CDCl_3$]: δ 2.75 (dd, 1H, *J* 2.16 Hz, 5.61 Hz, 2'-CH), 2.81 (dd, 1H, *J* 2.18 Hz, 6.61Hz, 2'-CH), 4.36 (m, 1H, 4'-CH), 4.41 (m, 2 H, 5'-CH_2), 5.22 (m, 1H, 3'-CH), 6.25 (t, 1H *J* 5.60, 1'-CH) 7.86 (s, 1H, 6-CH); ^{13}C NMR [$CDCl_3$]: δ 12.2 (5-CH_3), 20.7 (CH_3 (acetyl)), 20.8 (acetyl CH_3), 38.8 (2'-CH_2) 63.6 (5'-CH_2), 73.9 (3'-CH), 83.0 (1'-CH), 87.2 (4'-CH), 103.4 (5-C), 141.8 (C-4), 154.4 (PFP C), 168.3 (2-CO), 170.2 (acetyl CO), 170.4 (acetyl CO); ^{19}F NMR [$(CD_3)_2SO$] δ;-15.0 (t, 2 F, *J* 20.2 Hz, Ar-F), -19.9 (t, 1 F, *J* 21.2 Hz, Ar-F), -24.3 (d, 2 F, *J* 24.2 Hz, Ar-F); mass spectrum (FAB^+): *m/z* 1007 (2M+Na, 10%), 985 (2M+H, 18%), 515 (M+Na, 50%), 493 (M+H, 37%), 293 (100%), 201 (62%) [Found: *m/z* 493.1049 (M+H).$C_{20}H_{18}F_5N_2O_7$ requires 493.1034].

1-β-D-*Erythro-*pentofuranosyl-4-pentafluorophenoxy-pyrimidine-2(1H)-one 95: Uridine (2 g, 8.2 mmol) was dried by co-evaporation with anhydrous pyridine (3x20 mL), dissolved in pyridine (30 mL), placed under Ar and the solution was cooled to 0°C to which trifluoroacetic anhydride (5.80 mL, 41 mmol) was added and the reaction mixture stirred at room temperature for 24 h. Pentafluorophenol (20 g, 43.9 mmol) in pyridine (20 mL) was added and the reaction mixture was stirred for 72 h before being concentrated under vacuum. Flash chromatography, eluting with EtOAc-MeOH (4:1), yielded the title compound (1.98 g, 59%) as a white powder; m.p. 238°C; R_f 0.53 (EtOAc-MeOH 4:1); ^1H-NMR [$(CD_3)_2SO$]: δ 3.68 (m, 2 H, 5'-CH_2), 4.02 (m, 3 H, 2'-CH, 3'-CH, 4'-CH), 5.18 (s, 1 H, 2'-OH), 5.24 (s, 1 H, 3'-OH), 5.87 (s, 1 H, 5'-OH), 5.87 (s, 1 H, 1'-CH), 6.71 (d, 1 H, *J* 7.4 Hz, 5-CH), 8.66 (d, 1 H, *J* 7.4 Hz, 6-CH); ^{13}C-NMR (d_6-DMSO) 59.6 (5'-CH_2), 68.4 (2'-CH), 74.5 (3'-CH), 84.3 (1'-CH), 90.6 (4'-CH), 92.7 (5-CH), 147.9 (6-CH), 153.7 (4-C), 168.5 (2-CO); ^{19}F-NMR [$(CD_3)_2SO$]: δ -14.76 (t, 2 F, *J* 20.1 Hz, 2x*Meta* CF), -19.61 (t, 1 F, *J* 21.2 Hz, *Para* CF), -21.5 (d, 2 F, *J* 24.2 Hz, 2x*Ortho* CF); Mass Spectrum (FAB^+):

m/z: 411 (M+H, 70%), 308 (17%), 297 (100%), 156 (15%), 121 (19%), 108 (108%); Anal. Calcd. for $C_{15}H_{11}F_5N_2O_5$; C, 43.9; H, 2.6; N, 6.8. Found: C, 43.9; H, 2.6; N, 7.2.

1-β-D-*Erythro-*pentofuranosyl-4-(1,2,4-triazolyl)-pyrimidine-2(1H)-one 96: Uridine (1.01 g, 4.45 mmol) was dried by co-evaporation with anhydrous pyridine (3x40 mL), dissolved in pyridine (50 mL), placed under Ar and the solution cooled to 0°C. Trifluoroacetic anhydride (3.1 mL, 2.23 mmol) was added and the reaction mixture stirred at room temperature for 24 h. Pentafluorophenol (8.1 g, 43.90 mmol) in pyridine (10 mL) was then added and the mixture stirred for 72 h before being concentrated under vacuum. Flash chromatography, eluting with CH_2Cl_2-CH_3OH (4:1) followed by crystallisation from DMF, yielded the title compound (684 mg, 56%) as a colourless solid: m.p. 216°C; TLC (CH_2Cl_2-CH_3OH 4:1): R_f 0.45; UV (95% EtOH): λ_{max} 249.0 nm (8200), 315 nm (4900);IR (KBr disc): υ_{max} 3357, 3300, 3118, 3095, 2933, 1697, 1630, 1562, 1523, 1417, 1285, 1107, 1068 cm^{-1}; ^1H NMR [$(CD_3)_2SO$]: δ 4.32 (m, 1 H, 3'-CH), 4.34 (m, 1 H, 4'-CH), 4.53 (m, 3 H, 2'-CH, 5'-CH_2), 5.41 (d,1 H, *J* 6.0 Hz, 2'-OH), 5.72 (t, 1 H, *J* 5 Hz, 5'-OH), 6.10 (d, 1 H, *J* 4.5 Hz, 3'-OH), 6.31 (d, 1 H , *J* 1.4 Hz, 1'-CH), 7.43 (d, 1 H, *J* 7.3 Hz, 5-CH), 8.81 (s, 1 H, 3-CH (triazolyl)), 9.42 (d, 1 H, *J* 7.3 Hz, 6-CH), 9.85 (s, 1 H, 5-CH (triazolyl)); ^{13}C NMR [$(CD_3)_2SO$]: δ 53.2 (5-CH_2), 61.9 (2'-CH), 68.4 (3'-CH), 77.98 (1'-CH), 84.9 (4'-CH), 87.5 (5-CH), 137.53 (triazolyl 5-CH) 142.2 (6-CH), 147.6 (C-4), 147.9 (3-CH (triazolyl)), 152.4 (2-C); Mass Spectrum (FAB): *m/z* 318 (M+ Na, 22%), 296 (M+H, 37%), 242 (20%), 165 (86%), 150 (49%), 115 (81%), 105 (100%); Anal. Calcd for $C_{11}H_{13}N_5O_5$: C, 44.7; H, 4.4; N, 23.7. Found: C, 44.7; H, 4.3; N, 23.4.

1-(5'-O-Acetyl-3'-azido-2',3'-dideoxy-β-D-*erythro-*pentofuranosyl)-5-methyl-4-pentafluorophenoxy-pyrimidine-2(1H)-one 97: Nucleoside **85** (1.00 g, 3.24 mmol) was dried by co-evaporation with anhydrous pyridine (3x20 mL), dissolved in pyridine (30 mL), placed under Ar and the solution cooled to 0°C. 4-Chlorophenyl phosphorodichloridate (1.05 mL, 6.47 mmol) was added and the reaction mixture stirred at room temperature for 24 h. Pentafluorophenol (2.38 g, 12.9 mmol) was added and the reaction mixture stirred for

a further 72 h before being concentrated under vacuum. Flash chromatography, eluting with EtOAc-Hexane (1:1), gave the title compound as a gum (0.95 g, 62%); TLC (EtOAc-hexane 1:1): R_f 0.51; UV (95% EtOH) λ_{max} 296 nm (7220); IR (KBr disc): υ_{max} 3095, 2940, 2665, 2465, 2112, 1680, 1550, 1515, 1398, 1330, 1232, 1111, 992, 727 cm^{-1}; ^1H NMR [$CHCl_3$]: δ 2.17 (s ,3 H, CH_3 (acetyl)), 2.25 (s, 3 H, 5-CH_3), 2.40 (m, 1 H, 2'-CH), 2.73 (m, 1 H, 2'-CH), 4.16 (m, 2 H, 5-CH_2), 4.42 (m, 2 H, 3'-CH, 4'-CH), 6.09 (dd, 1 H, *J* 4.8 Hz, 1.7 Hz, 1'-CH), 7.86 (s, 1 H, 6-CH); ^{13}C NMR [$(CD_3)_2SO$]: δ 12.3 (5-CH_3), 20.8 (CH_3 (acetyl)), 38.6 (2'-CH_2), 59.8 (3'-CH), 62.9 (5'-CH_2), 82.4 (1'-CH), 87.3 (4'-CH), 103.1 (5-C), 142.1 (6-C), 154.2 (4-C), 167.5 (2-C), 170.1 (CO(acetyl)); Mass Spectrum (FAB^+): *m/z* 973 (2M+Na, 19%), 951 (2M+H, 35%), 498 (M+Na, 56%), 293 (100%), 184 (31%) [Found: *m/z* 476.1002 (M+H).$C_{18}H_{15}F_5N_5O_5$ requires 476.0993]; Anal. Calcd. for $C_{18}H_{14}F_5N_5O_5$; C, 45.5; H, 2.9; N, 14.7. Found: C, 45.7; H, 2.7; N, 14.3.

1-(5'-O-Acetyl-3'-azido-2',3'-dideoxy-β-D-*erythro*-pentofuranosyl)-5-methyl-4-(1,2,4-triazolyl)-pyrimidine-2(1H)-one 98: Nucleoside **85** (3.08 g, 9.9 mmol) was dried by co-evaporation with anhydrous pyridine (3x30mL), dissolved in pyridine (50 mL), placed under Ar and the solution cooled to 0°C. 4-Chlorophenyl phosphorodichloridate (1.85 mL, 11.34 mmol) was added and the mixture stirred for 5 minutes. 1,2,4-1*H*-Triazole (2.7 g, 39.4 mmol) was added and the mixture stirred at room temperature for a further 72 h before being concentrated under vacuum. Flash chromatography, eluting with EtOAc, followed by crystallisation from EtOAc, give the title compound as colourless crystals (1.79 g, 88%); m.p. 137 °C; TLC (EtOAc): R_f 0.24; UV(95% EtOH): λ_{max} 248 (5500), 327 (3900); IR (KBr disc): υ_{max} 3126, 2970, 2112, 1730, 1662, 1625, 1525, 1425, 1237, 977 cm^{-1}; ^1H NMR [$(CD_3)_2SO$]: δ 2.09 (s, 3 H, CH_3 (acetyl)), 2.35 (s, 3 H, 5-CH_3), 2.59 (m, 2 H, 2'-CH_2), 4.19 (m, 1 H, 3-CH), 4.39 (m, 3 H, 4'-CH, 5'-CH_2), 6.12 (t, 1 H, *J* 5.6 Hz, 1'-CH), 8.23 (s, 1 H, 6-CH), 8.40 (d, 1 H, 3-CH (triazolyl)), 9.35 (s, 1 H, 5-CH (triazolyl)); ^{13}C NMR [$(CD_3)_2SO$]: δ 21.2 (CH_3 (acetyl)), 25.7 (5-CH_3), 42.3 (2'-CH_2), 64.7 (5'-CH), 68.1 (5'-CH_2), 87.1 (1'-CH), 92.0 (4'-CH), 109.8 (5-CH), 150.4 (6-CH), 150.8 (5-CH (triazolyl)), 158.0 (4-C), 158.6 (3-CH (triazolyl)), 163.1 (2-C), 175.3 (CO (acetyl)); *m/z* (FAB^+) 383

(M+Na, 1%), 361 (M+H, 37%), 178 (100%), 135 (5%); Anal. Calcd for $C_{14}H_{16}N_8O_4$: C, 46.7; H, 4.4; N, 30.8. Found: C, 46.7; H, 4.4; N, 31.1.

Attempted synthesis of 1-(3'-Azido-2',3'-dideoxy-β-D-*erythro*-pentofuranosyl)-5-methyl-4-pentafluorophenoxy-pyrimidine-2(1H)-one 99: Method A. AZT (500 mg, 1.87 mmol) was dried by co-evaporation in pyridine (3x20 mL), dissolved in pyridine (30 mL), placed under Ar and the solution cooled to 0°C to which trifluoroacetic anhydride (0.79 mL, 5.62 mmol) was added and the reaction mixture stirred at room temperature for 24 h. Pentafluorophenol (2.41 g, 13.1 mmol) was added and the reaction mixture stirred for a further 72 h at room temperature before being concentration under vacuum to give a dark brown syrup. TLC indicated mainly starting material was present.

1-(3'-Azido-2',3'-dideoxy-β-D-*erythro*-pentofuranosyl)-5-methyl-4-pentafluorophenoxy-pyrimidine-2(1H)-one 99: Method B. AZT (500 mg 1.87 mmol) was dried by co-evaporation with anhydrous pyridine (3x10 mL), dissolved in pyridine (10 mL) and placed under Ar to which trimethylsilyl chloride (0.48 mL, 3.74 mmol) was added and the reaction mixture stirred at room temperature for 30 minutes before being cooled to 0°C. 4-Chlorophenyl phosphorodichloridate (0.76 mL, 4.68 mmol) was added and stirred for 15 minutes followed by pentafluorophenol (1.73 g, 9.36 mmol, 5) in pyridine (10 mL) and the reaction mixture stirred at room temperature for 72 h. Water (50 mL) was then added and the mixture stirred for a 1 h at room temperature before being concentrated under vacuum. Flash chromatography, eluting with EtOAc-hexane (7:1), yielded the title compound (502 mg, 62%) as a foam; m.p. 181°C; TLC (EtOAc-Hexane 7:1): R_f 0.61; UV (95% EtOH) λ_{max} 295 nm (9820); IR (KBr disc): ν_{max} 3449, 2931, 2669, 2362, 2117, 2092, 1664, 1521, 1411, 1330, 1112, 988, 778 cm^{-1}; ^{1}H-NMR [$(CD_3)_2SO$]: δ 2.13 (s, 3 H, 5-CH_3), 2.42 (t, 2 H, J 6.4 Hz, 2'-CH_2), 3.70 (m, 1 H, 5'-CH), 3.91 (m, 1 H, 4'-CH), 4.40 (q, 1 H, J 6.7 Hz, 3'-CH), 5.40 (s, 1 H, 5'-OH), 6.61 (t, 1 H, J 4.3 Hz, 1'-CH), 8.42 (s, 1 H, 6-CH); ^{13}C-NMR [$(CD_3)_2SO$]: δ 1.4 (5-CH_3), 28.5 (2'-CH_2), 50.8 (3'-CH), 51.0 (5'-CH_2), 75.9 (1'-CH), 76.9 (4'-CH), 92.3 (5-C), 133.9 (6-CH), 144.3 (4-C), 158.5 (2-C); ^{19}F-NMR [$(CD_3)_2SO$] δ

-14.83 (t, 2 F, *J* 22.3 Hz, 2x*Meta* CF), -19.61(t, 1 F, *J* 23.2 Hz, *Para* CF), (d, 2 F, *J* 23.6 Hz, 2x*Para* CF); Mass Spectrum (FAB^+): *m/z* 434 (M+H, 27%), 319 (10%), 293 (100%), 165 (42%), 128 (37%), 115 (34%), 105 (58%) [Found: *m/z* 434.0873 (M+H).$C_{16}H_{12}F_5N_5O_4$ requires 440.0887]; Anal. Calcd. for $C_{16}H_{12}F_5N_5O_4$: C, 44.3; H, 2.8; N, 16.2. Found: C, 44.6; H, 2.7; N, 16.1.

Attempted synthesis of 1-(3'-azido-2',3'-dideoxy-β-D-*erythro*-pentofuranosyl)-5-methyl-4-(1,2,4-triazolyl)-pyrimidine-2(1H)-one 100: Method A. AZT (1.00 g, 3.75 mmol) was dried by co-evaporation in pyridine (3x30 mL), dissolved in pyridine (50 mL), placed under Ar and the solution cooled to 0°C. Trifluoroacetic anhydride (1.59 mL, 11.2 mmol) was added and the reaction mixture stirred at room temperature for 24 h before 1,2,4-1*H*-Triazole (1.55 g, 22.4 mmol) was added and the reaction mixture stirred at room temperature for 72 h before being concentrated under vacuum to give a dark brown syrup. TLC analysis indicated mainly starting material was present and did not show the characteristic fluorescent triazolyl-nucleoside spot that would be expected under UV light.

Attempted synthesis of 1-(3'-azido-2',3'-dideoxy-β-D-*erythro*-pentofuranosyl)-5-methyl-4-(1,2,4-triazolyl)-pyrimidine-2(1H)-one 100: Method B. AZT (1.00 g, 3.75 mmol) was dried by co-evaporation in pyridine (3x30 mL) dissolves in pyridine (50 mL) and placed under Ar to which trimethylsilylchloride (0.52 mL, 4.12 mmol) was added and the reaction mixture stirred for 4 h before being cooled to 0°C. 4-Chlorophenyl phosphorodichloridate (1.34 mL, 8.23 mmol) was added and the reaction mixture stirred at room temperature for 24 h to which 1,2,4-1*H*-triazole (1.03 g, 15.0 mmol) was added and the reaction mixture stirred for 72 h. Water (10 mL) was added and stirred for 1 minute before being concentrated under vacuum to give a dark brown syrup. TLC analysis indicated a number UV active spots including starting material. A spot was seen under UV, characteristic of a triazolyl nucleoside, but any attempt to isolated the title compound by flash chromatography failed.

6.2 Oligonucleotide Analysis

6.2.1 General methods

HPLC analyses and purification's were carried out on a Water's 600 Multisolvent delivery system with UV detection at 260 nm and chromatograms recorded on a Hewlett-Packard 3396 integrator. Injection volumes were 50 μL for an analytical run and 0.5 mL for a semi-prepative run.

6.2.2 HPLC analysis

Mobile phase for gradient HPLC analysis and purification consisted of an aqueous CH_3CN gradient ranging from 2% to 35% acetonitrile during 20 minutes buffered with 0.1 M triethylammonium acetate (Section 6.2.4). Isolation of the oligonucleotide required de-salting of the triethylammonium buffer. Triethylammonium acetate is a volatile salt and was removed by co-evaporation as described below:

i. evaporation of the mobile phase was performed under vacuum;

ii. the residue from **i** was suspended in water (1 mL) and evaporated under vacuum;

iii. the residue from **ii** was suspended in 50% ethanol solution (1 mL) and evaporated under vacuum.

6.2.3 Preparation of 1.0 M triethylammonium acetate buffer stock solution:

HPLC grade acetic acid	28.6 mL
HPLC grade triethylamine	69.7 mL
HPLC grade water	to 500 mL.

To about 300 mL of water, triethylamine (69.7 mL) was added and the mixture stirred. Acetic acid (28 mL) was then added slowly. Water was added to increase to total volume of solution to about 490 mL. The pH of the buffer was monitored using a pH meter whilst acetic acid was added drop-wise until pH 7.0 was achieved. Water was then added to make a total volume of 500 mL.

6.2.4 Preparation of 2% CH_3CN in 0.1 M triethylammonium acetate buffer (Buffer A):

To about 200 mL of water, 1.0 M triethylammonium acetate buffer stock solution (50 mL) was added. To this CH_3CN (10 mL) was added and the solution was made up to a total volume of 500 mL with HPLC grade water.

6.2.5 Preparation of 80% CH_3CN in 0.1 M triethylammonium acetate buffer (Buffer B):

To 100 mL of water, 1.0 M triethylammonium acetate buffer stock solution (50 mL) was added. The solution was made up to a total volume of 500 mL with CH_3CN.

6.2.6 HPLC gradient:

The use of gradient HPLC allows improved resolution of oligonucleotides which may only differ by one base-pair. HPLC analyses and purification's of oligonucleotides were carried out using an CH_3CN gradient. The gradient ran from 0% to 50% buffer B in buffer A during 25 minutes (Figure 6.1). The gradient continued to 100% buffer B during 5 minutes to ensure the HPLC column was clear of any hydrophobic residues before returning to 100% buffer A in preparation for the run (Figure 6.1). Note: A period of at least 10 minutes was necessary to allow the HPLC column to re-equilibrate.

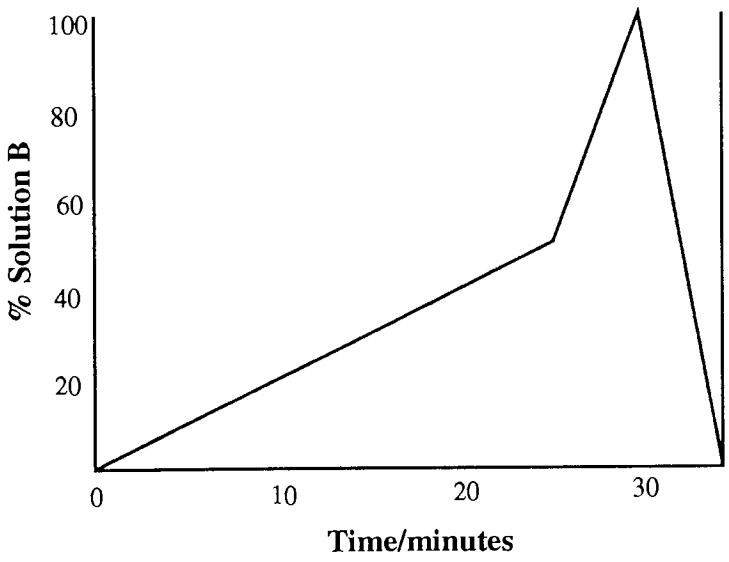


Figure 6.1 Schematic representation of HPLC gradient used to analyse and purify oligonucleotides.

6.3 Thermal analysis of antisense oligonucleotides

6.3.1 General methods

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 of -10°C to >100°C.

6.3.2 UV thermal analysis

Oligonucleotides for thermal analysis were dissolved in 1.0 M aqueous sodium chloride buffer, placed in a 1 cm UV cuvette and cooled to -2°C from 80°C at a rate of 0.1°C per minute, to aid annealing. Condensation and frosting were prevented by passing a dry stream nitrogen over the UV cuvettes. Oligonucleotide samples were heated at a rate of 0.5°C per minute from 0°C to 80°C; the absorbance against temperature readings were recorded automatically at 0.5°C intervals by Cary Thermal Analysis software using a Victor 433D personal computer.

6.3.3 Preparation of 2.0 M NaCl UV stock solution for thermal analysis:

NaCl	11.6 g
TRIS (hydroxymethyl)aminomethane	0.24 g
0.01 M TRIS(hydroxymethyl)aminomethane.HCl	to pH 7
HPLC grade water	to 100 mL

To about 70 mL of water NaCl (11.6 g) and TRIS(hydroxymethyl)aminomethane (0.24 g) was added and the solution stirred. Water was added to increase the volume to 95 mL to which 0.05 M TRIS(hydroxymethyl)aminomethane.HCl solution was added drop-wise until pH 7 had been achieved.

6.3.4 Preparation of 1.0 M NaCl solution for UV thermal analysis buffer:

To a dried sample of oligonucleotide in an Ependorf, 2.0 M NaCl UV thermal analysis stock buffer (0.65 mL) and water (0.65 mL), were added. The Ependorf was shaken overnight before UV thermal analysis.

References

1. Feinberg, M. B. *Lancet* **1996**, *348*, 239-246.
2. Clerici, M.; Levin, J. M.; Keessler, H. A.; Harris. H. A.; Berzofsky, J. A.; Landay, A. L.; Shearer, G. M. *J. Am. Med. Assoc.* **1994**, *271*, 42-46.
3. Mertens, T. E.; Belsey, E.; Stoneburner, R. L.; Beer, D. L.; Sato, P.; Burton, A.; Merson, M. H. *AIDS* **1995**, *9*, S259-S272.
4. Bhave, G.; Lindan, C. P.; Hudes, E. S.; Desai, S.; Wagle, U.; Tripathi, S.P.; Mandel, J. S. *AIDS* **1995**, *9*, S21-S30.
5. Choi, K -H.; Lew, S.; Vittinghoff, E.; Catania J. A.; Barrett, D. C.; Coaes, T. J. *AIDS* **1996**, *10*, 81-87.
6. Fenyo, E. M. *Immunol. Rev.* **1994**, *140*, 131-146.
7. Gelderblom, H. R.; Ozel, M.; Pauli, G. *Arch.Virol.* **1989**, *106*, 1-13.
8. Gelderblom, H. R.; Hausmann, E. H. S.; Ozel, M.; Pauli, G.; Knock, M. A. *Virology* **1987**, *156*, 171-176.
9. Boshl, M. L.; Earl, P. A.; Fargnoli, K.; Picciafuoco, S.; Giombini, F. *Science* **1989**, *244*, 694-697.
10. Leis, J.; Baltimore, D.; Bishop, J. M.; Coffin, J.; Fleissner, E.; Goff, S. P.; Oroszlan, S.; Robinson, H.; Skalka, S. M.; Temin, H. M.; Vogt, V. *J. Virol.* **1988**, *62*, 1808-1809.
11. Dayton, A. I.; Sodroski, J. G.; Rosen, C. A.; Goh, W. C.; Heseltine, W. A. *Cell* **1986**, *44*, 941-947
12. Terwilliger, E. F.; Cohen, E. A.; Lu, Y.; Sodroski, J. G.; Heseltine, W. A. *Proc. Natl. Acad. Sci. USA.* **1989**, *86*, 5163-5167.
13. Gowda, S. D.; Stein, B. S.; Engleman, E. G. *J. Biol. Chem.* **1989**, *264*, 8459-8462.
14. McKeating, J. A.; Willey, R. L. *AIDS* **1989**, *3*, S35-S41.
15. Kohl, N.E.; Emini, E. A.; Schleif, W. A.; Davis, L. J.; Heimbeck, J. C.; Dixon, R. A. F.; Scolnic, E. M.; Sigel, I. S. *Natl. Acad. Sci. USA.* **1988**, *85*, 4686-4690.

16. Veronese, F. D.; Copelamd, T. D.; DeVico, A. L.; Rahman, R.; Oroszlan, S.; Gallo, R. C.; Sarngadharam, M. G. *AIDS Res. Hum. Retroviruses* **1987**, *3*, 253-264.

17. Humber, H. E.; McCoy, J. M.; Seehra, J. S.; Richardson, C. C. *J. Biol. Chem.* **1989**, *264*, 4669-4678.

18. Oyama, F.; Kikucki, R.; Crouch, R. J.; Uchida, T. *J. Biol. Chem.* **1989**, *264*, 18808-18817.

19. Haseltine, W. A. *AIDS* **1989**, *2*, 311-334.

20. Cheynier, R.; Henrickwark, S.; Hadida, F.; Pelletier, E.; Oskenhendled, H.; Autran, B.; Wain-Hubson, S. *Cell* **1994**, *78*, 373-387.

21. Barter, G.; Barton, S.; Gazzard, B. *HIV and AIDS*; **1993**, Churchill Livingstone: Edinburgh.

22. Bolognesi, D. P. *Adv. Viral Res.* **1993**, *42*, 103-148.

23. Geleziunas, R.; Bour, S.; Wainberg, M. A. *Adv. Viral Res.* **1994**, *44*, 203-265.

24. Pantaleo, G.; Graziosi, C.; Fauci, A. S.; *New Engl. Med.* **1993**, *328*, 327-335.

25. Coombs, R. W.; Collier, A. C.; Allain, J. P.; Nihora, B.; Leither, M.; Gjerst, G. F.; Corey, L. *N. Engl. J. Med.* **1989**, *321*, 1626-1631.

26. Fauci, A. S.; Schnittman, S. M.; Poli, G.; Koening, S.; Pantaleo, G. *Ann. Intern. Med.* **1991**, *114*, 678-693.

27. Rosenberg, Z. F.; Fauci, A. S. *The FABEB Journal.* **1991**, *5*, 2382-2390.

28. Ho, D. D.; Newman, A. U.; Perelson, A. S.; Chen, W.; Lenard, J. M.; Markowitz, M. *Nature* **1995**, *373*, 123-126.

29. Wei, X.; Ghosh, S. K.; Taylor, M. E.; Johnson, V. A.; Emini, E. A.; Deutsch, P.; Lifson, J. D.; Bonhoeffer, S.; Nowak, M. A.; Hahn, B. H.; Saag, M. S.; Shaw, G. M. *Nature* **1995**, *373*, 117-122.

30. Clerici, M.; Shear, G. M. *Immunol. Today* **1993**, *14*, 107-111.

31. Clerici, M.; Haklim, F. T.; Venzon, D. J.; Blatt, S.; Hendrix, C. W.; Wynn, T. A.; Shearer, G. M. *J. Clin. Invest.* **1993**, *91*, 759-760.

32. Grant, M. D.; Smaill, M. F.; Rosenthal, K, L. *J. AIDS.* **1994**, *7*, 571-579.

33. Zinkernagel, R. M.; Hengartner, *H. Immunol. Today* **1994**, *15*, 262-268.

34. Langtry, H. D.; Camoli-Richards, D. M. *Drugs* **1989**, *15*, 262-268.

35. Faulds, D.; Brogen, R. N.; *Drugs* **1992**, *44*, 94-116

36. Whittington, R.; Brogdin, R. N. *Drugs* **1992**, *44*, 656-683.

37. Hitchcock, M. J. M. *Antiviral Chemisty and Chemotherapy* **1991**, *2*, 125-132.

38. Horwitz, J. P.; VChau, J.; Noel, M. *J. Org. Chem.* **1964**, *29*, 2076-2078.

39. Ostertag, W.; Roesler, G.; Krieg, C. J.; Kind, J.; Cole, T.; Crozier, T.; Gaedicke, G.; Steinheider, G.; Kluge, N.; Dube, S. *Proc. Natl. Acad. Sci. USA.* **1974**, *71*, 4980-4985.

40. Mitsuya, H.; Weinhold, J. J.; Furman, P. A.; St Clair, M. H.; Lehrman, S. W.; Gallo, R. C.; Boiognesi, D.; Barry, D. W.; Broder, S. *Proc. Natl. Acad. Sci. USA.* **1985**, *82*, 7096-7100.

41. Soudeyns, H.; Yao, X-J.; Gao, Q.; Belleau, B. Kraus, J. -L.; Nguyen-Ba, N.; Spira, B.; Wainberg, M. A. *Antimicrob. Agents Chemother.* **1991**, *35*, 1386-1390.

42. Sommadosi, J. P.; Schinazi, R. F.; Chin, C. K.; Xie, M. Y. *Biochem. Pharmacol.* **1992**, *44*, 1921-1925.

43. Pauwels, R.; Andries, K.; Desmyter, J.; Kukla, M.; Breslin, H. J.; Raeymaeckers, A.; Van Gelder, J.; Woestenborghs, R.; Haykants, J.; Schellekens, K; Janssen M. A. C.; De Clercq, E.; Janssen, P. A. J. *Nature* **1990**, *343*, 470-474.

44. Debyser, Z.; Pauwel, R.; Andries, K.; Desmyter, J.; Kukla, M.; Janssen, P. A. J.; De Clercq, E. *Proc. Natl. Acad. Sci. USA.* **1991**, *88*, 1451-1455.

45. AmFar *AIDS/HIV Treatment Directors* **1992**, *40.*

46. Merluzzi, V. K.; Hargrave, K. D.; Labadia, M.; Grozinger, K.; Skog, M.; Skoog, M.; Wu, J. C.; Shih, C.; Eckner, K.; Hattox, S.; Adams, J.; Rosehthal, A. S.; Faanes, R.; Eckner, R. J.; Koop, R. A.; Sullivan, J. L.; *Science* **1990**, *250*, 1411-1413.

47. Richmond, D. D. *AIDS. Res. & Hum. Retroviruses* **1992**, *8*, 1065-1071.

48. Kohl, N. E.; Emini, E. A.; Scielf, W. A.; Davis, L. J.; Heimbach, J. C.; Dixon, R. A. F; Scolnick, E. M.; Sigal, I. S. *Proc. Natl. Acad. Sci. USA.* **1988**, *85*, 4686-4690.

49. Roberts, N. A.; Martin, J. A.; Kinchington, D.; Broadhurst, A. V.; Craigh, J. C.; Duncan, I. B.; Galpin, S. A.; Handa, B. K.; Kay, J.; Korhn, A.; Lambert, R. W.;

Merret, J. H.; Mills, J. S.; Parkes, K. E. B.; Redshaw, S.; Richie, A. J.; Taylor, D. L.; Thomas, G. J.; Machin, P. J. *Science* **1990**, *248*, 385-361.

50. Winslow, D. L.; Otto, M. *J. AIDS.* **1995**, *9*, 5183-5192.

51. Jacobsen, H.; Yasargil, K.; Winslow, D. L.; Craig, J. C; Krohn, A.; Duncan, I. B.; Mous, J. *Virology* **1995**, *206*, 527-534.

52. Dorsey, B. D.; Levin, R. B.; McDaniel, S. L.; Vacca, J. P; Guare, J. P.; Darke, P. L.; Zaugay, J. A.; Emini, E. A.; Schleif, W. A.; Quintero, J. C.; Lin, J. H.; Chen, I. W.; Holloway, M. K.; Fitzgerald, M. K.; Fitzgerald, P. M. D.; Axel, M. G.; Ostovic, D.; Anderson, P. S.; Joel, R. H. *J. Med. Chem.* **1994**, *37*, 3443-3451.

53. Vacca, J. P.; Dorsey, B. D.; Scheif, W. A.; Levin, R. B.; McDaniel, S. L.; Drake, P. L.; Zugay, J.; Quintero, J. C.; Blahy, O. M.; Roth, E.; Sardana, V. V.; Schlabach, A. J.; Graham, P.I; Condra, J. H.; Gotlib, L.; Holloway, M. K.; Lin, J.; Chem, I.-W.; Vastag, K.; Ostovic, D.; Anderson, P. S.; Emini, E. A.; Huff, J. R. *Proc. Natl. Acad. Sci. USA.* **1994**, *91*, 4096-4100.

54. Johnson, V. A. *AIDS. Res. & Hum. Retroviruses* **1994**, *10*, 907-912.

55. Caliendo, A. M.; Hirsch, M. S. *Clin. Infect. Dis.* **1994**, *18*, 516-524.

56. Yarchoan, R, Lietzau, J. A.; Nguyen, B-Y. *J. Infect. Dis.* **1994**, *169*, 9-17.

57. Fischl, M. A.; Stanley, K.; Collier, A. C.; Arduino, J. M.; Stein, D. S.; Feinberg, J. E.; Allen, D.; Goldsmith, J. C. ; Powderly, W. G. *Ann. Intern. Med.* **1995**, *122*, 24-32.

58. de Jong, M. D.; Loewenthal, M.; Boucher, C. A. B.; Van der Ende, I.; Hall, D.; Schipper, P.; Imrie, A.; Weigel, H. M.; Kauffmann, R. H.; Koster, R.; Seville, P.; Rocklin, R.; Cooper, D. A.; Lange, J. M. A. *J. Infect. Dis.* **1994**, *169*, 1346-1350.

59. Collier, A. C.; Coombs, R. W.; Timpone, J.; van der Ende, I.; Hall, D. Schipper. P.; Imrie, A.; Weigel, H. M.; Kauffmann, R, H.; Kester, R.; Seville, P.; Rocklin, R.; Cooper, D. A.; Lange, J. M. A. *Proceedings of the 10th conference on AIDS*, Yokohoma, Japan, Abstract 058B, **1994**.

60. Burke, D. S. *AIDS* **1995**, *Supplement A*, S171-S180.

61. Cohen, I. *Science* **1993**, *262*, 980-998.

62. Matthews, T. *J. AIDS Res. Hum. Retroviruses* **1994**, *10*, 631-632.

63. Poeschla, E. M.; Wong-Staal, F. *AIDS Clin. Rev.* **1995**, *196*, 1-45.

64. Pleaner, P.; Goodchild, J.; Kalckar, H. M; Zamecnick, P. C. *Proc. Natl. Acad, Sci. USA* **1987**, *84*, 1936-1939.

65. *Oligonucleotide Synthesis-A Pratical Approach*; Gait, M. L. Ed.; IRL Press: Oxford, 1984.

66. Watson, J. D.; Crick, F. H. *Nature* **1953**, *171*, 737-738.

67. Uhilmann, E.; Peyman, A. *Chem. Rev.* **1990**, *90*, 544-584.

68. Beaucage, S. L.; Iyer, R. P. *Tetrahedron* **1993**, *49*, 6123-6194.

69. Hurley, L. H.; Boyd, F. L. *Trends Pharmacol. Sci.* **1988**, *9*, 402-407.

70. Zamecnik, P. C.; Goodchild, J.; Taguchi, Y.; Sarin, P. S. *Proc. Natl. Acad. Sci. USA.* **1986**, *83*, 4143-4146.

71. Matsuhura, M.; Shinozuka, K.; Zon, G.; Mitsuya, H.; Reitz, M.; Cohen, J. S.; Broder, S. *Proc. Natl. Acad. Sci. USA.* **1987**, *84*, 7706-7710.

72. Goodchild, J.; Agrawal, S.; Civeira, M. P.; Sarin, P. S.; Sun, D.; Zamecnik, P. C.; *Proc. Natl. Acad. Sci. USA.* **1988**, *85*, 5507-5511.

73. Sarin, P. S.; Agrawal, S.; Civeria, M. P.; Goodchild, J.; Ikeuchi, T.; Zamecnik, P. C. *Proc. Natl. Acad. Sci. USA.* **1988**, *85*, 7448-7451.

74. Zaia, J. A.; Rossi, J. J.; Marakawa, G. J.; Spallone, P. A.; Stephens, D. A.; Kaplan, B. E.; Eritja, R.; Wallace, R. B.; Cantin, E. M. *J. Virol.* **1988**, *62*, 3914-3917.

75. Agrawal, S.; Ikeuchi, T.; Sun, D.; Sarin, P. S.; Konopka, A.; Maizel, J.; Zamecnik, P. C. *Proc. Natl. Acad. Sci. USA.* **1989**, *86*, 7790-7795.

76. Matsukura, M.; Zon, G.; Shinozuka, K.; Robert-Gurot, F. M.; Shimada, T.; Stein, C. A.; Mitsuya, H.; Wong-Staal, F.; Cohern, J. S.; Broder, S. *Proc. Natl. Acad. Sci. USA.* **1989**, *86*, 4244-4248.

77. Daum, T.; Engele, J. W.; Mag, M.; Muth, J.; Lucking, M.; Schroder, H. C.; Matthews, E.; Muller, W. E. *Intervirology* **1992**, *33*, 65-75.

78. Lisziewicz, J.; Sun, D.; Koltman, M.; Agrawal, S.; Zamecnik, P. C.; Gallo, R. *Proc. Natl. Acad. Sci. USA.* **1992**, *89*, 11209-11213.

79. Rhodes, A.; James, W. *J. Gen. Virol* **1990**, *71*, 1965-1974.

80. Sczakie, G.; Pawlita, M. *J. Virol.* **1991**, *65*, 468-472.

81. Joshi, S.; Van Brunscroft, A.; Asad, S.; van der Elst, I.; Read, S. E.; Burnstein, A. *J. Virol.* **1991**, *65*, 5524-5530.

82. Rhodes, A.; James, W. *AIDS* **1991**, *5*, 145-151.

83. Whitton, J. L. *Advan. Virus Res.* **1994**, *44*, 267-303.

84. Ojwang, J. O.; Hampel, A.; Looney, D. J.; Wong-Staal, F.; Rappaport, J. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 10802-10806.

85. Brown, T.; Hunter, W. N.; Leonard, G. A. *Chemistry in Britain* **1993**, *29*, 484-488.

86. *Principles of Nucleic Acid Structure;* Saenger, W. Ed.; Springer-Verlag: New York, 1984.

87. Blackburn, G. M.; Gait, M. J. *Nucleic Acids in Chemistry and Biology* **1990,** IRL Press: Oxford.

88. Dickerson, R. E.; Drew, H. R.; Connor, N. B.; Wing, R. M.; Fratini, A. V.; Kapka, M. L. *Science* **1982**, *216*, 475-485.

89. Berman, H. M. *Current Opinion in Structural Biology* **1991**, *1*, 423-427.

90. Kumar, V. D.; Weber, T. I. *Nucleic Acids Res.* **1993**, *21*, 2201-2208.

91. Kopka, M. L.; Fratini, A. V.; Drew, H. R.; Dickerson, R. E. *J. Mol. Biol.* **1983**, *163*, 129-146.

92. Prieve, G. G.; Yanagi, K.; Dickerson, R.E. *J. Mol. Biol.* **1990**, *217*, 177-199.

93. Norberg, J.; Nilsson, L. *J. Am. Chem. Soc.* **1995**, *117*, 10832-10840.

94. Breslauer, K. J., Frank, R., Blocker, H., Marky, L. A. *Proc. Natl. Acad. Sci. USA.* **1986**, *83*, 3746-3750.

95. Freier , S. M.; Kierzek, R.; Jaeger, J. A.; Sugimoto, N.; Caruthers, M. H.; Neilson, T. *Proc. Natl. Acad. Sci. USA.* **1983**, *85*, 9373-9377.

96. Cheng, Y.; Pettitt, B. M. *Prog. Biophys. Molec. Biol.* **1992**, *58*, 225-257.

97. M. S. Searle, D. H. Williams, *Nucleic Acids Res.* **1993**, *21*, 2051-2056.

98. Hunter, C. A.; Sanders, J. K. M. *J. Am. Chem. Soc.* **1990**, *112*, 5525-5534.

99. Maltseva, T. V.; Agback, P.; Chattopadhyaya, *J. Nucleic Acids Res.* **1993**, *21*, 4246-4252.

100. Tunis, M. J.; Hearst, J. E. *Biopolymers* **1968**, *6*, 1345-1353.

101. Anderson, C. F.; Record, M. T. *Annu. Rev. Biophys. Chem.* **1990**, *29*, 1253-1261.

102. Record, M. T.; Mazur, S. J.; Melancon, P.; Roe, J. H.; Shaner, S. L.; Unger, L. *Annu. Rev. Biochem.* **1981**, *50*, 997-1024.

103. Crooke, S. T. *FASEB Journal* **1993**, *7*, 533-539

104. Ilich, P.; Prendergast, F. G. *Biopolymers* **1992**, *32*, 667-694.

105. *Molecular biology of the cell*; Abots, B.; Bray, D.; Lewis, J.; Raff, M.; Roberts, K; Watson, J. D.; Garland Publishing: New York, 1993.

106. Hauptle, M, -T.; Frank, R.; Dobberstein, B. *Nucleic Acids Res.* **1986**, *14*, 1427-1448.

107. Stein, H.; Hausen, P. *Science* **1969**, *166*, 393-396.

108. Minshull, J.; Hunt, J. *Nucleic Acids Res.* **1986**, *14*, 6433-6451

109. Crum, C.; Johnson, J. D.; Nelson, A.; Roth, D; *Nucleic Acids Res.* **1988**, *16*, 4569-4581.

110. Walder, R. Y.; Walder, Y. A. *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 5011-5015.

111. Sproat, B. S.; Lamond, A. L.; Beijer, B.; Neuner, P.; Ryder, U. *Nucleic Acid Res.* **1989**, *17*, 3373-3386.

112. Bloch, E.; Lavigton, M.; Bertrand, J. R.; Pognan, F.; Morvan, F.; Malvy, C.; Rayner, B.; Imbach, J. L.; Paoletti, C. *Gene*, **1988**, *72*, 349-360.

113. Crooke, S. T. *Annu. Ref. Pharmacol. Toxicol.* **1992**, *32*, 329-376.

114. Maher, L. J.; Dolnick, B. J. *Nucleic Acids Res.* **1988**, *16*, 3341-3358.

115. Pitsch, S.; Wendeborn, S.; Juan, J.; Eschenmoser, A. *Helv. Chim. Acta.* **1993**, *76*, 2161-2183.

116. Groebke, K.; Fraser, W.; Hunziker, J.; Peng, L.; Diederichsen, U.; Zimmermann, K.; Holzner, A.; Delgado, G.; Leumann, C.; Eschenmoser, A. in preparation.

117. Engal, J. D.; von Hippel, P. H. *J. Biol. Chem.* **1978**, *253*, 927-932.

118. Delort, A. M.; Guy, A.; Molko, R.; Teoule, R. *Nucleosides and Nucleotides* **1985**, *4*, 201-203.

119. Sung, W. L. *J. Chem. Soc. Chem. Comm.* **1981**, 1089.

120. Sung, W. L. *Nucleic Acids Res.* **1981**, *9*, 6139-6151.

121. Gaffney, B. L.; Jones, R. A. *Tetrahedron Lett.* **1982**, *23*, 2253-2256.

122. Zhou, X, -X.; Welch, C. J.; Chattopadhyaya, J. *Acta Chemica Scandinavia* **1986**, *B 40*, 806-816

123. Zhou, X. -X.; Chattapadhyaya, J. *Tetrahedron* **1986**, *42*, 5149-5156

124. Gao, H.; Fathi, R.; Gaffney, B. L.; Goswami, B.; Kung, P.P.; Rhess, Y.; Jin, R.; Jones, R. A. *J. Org. Chem.* **1992**, *57*, 6954-6959.

125. Kamaike, K.; Takahashi, M.; Utsugi, K.; Tomizuka, K.; Ishido, Y. *Tetrahedron Lett.* **1995**, *36*, 3659-3762.

126. Kellenbach, E. R.; van den Elst, H.; Boelens, R.; van der Marel, G. A.; van Boom, J. H.; Kaptein, R. *Recl. Trav. Chim. Pays-Bas* **1991**, *110*, 387-388.

127. Cowart, M; Gibson, K. J.; Allen, D. J.; Benkovic, S. J. *Biochemistry* **1989**, *28*, 1975-1983.

128. Still, W. C.; Kahn, M.; Abhijit, M. *J. Org. Chem.* **1978**, *43*, 2923-2925.

129. Adamiak, R. W.; Biata, E. *Nucleic Acids Res.* **1985**, *13*, 2989-3003.

130. Bleasdale, C.; Ellwood, S. B.; Golding, B. T. *J. Chem. Soc. Perkin Trans. 1* **1990**, 803-804.

131. Webb, T. R.; Matteucci, M. D. *J. Am. Chem. Soc.* **1986**, *108*, 2764-2765.

132. *Introduction to Organic Chemistry;* Steitwieser, A.; Heathcock, C. H.; Macmillan Publishing Company: New York, 1985.

133. Koh, J. S.; Dervan, P. B. *J. Am. Chem. Soc.* **1992**, *114*, 1470-1478.

134. Piccirilli, J. A.; Krauch, T.; MacPherson, L. J.; Benner, S. A. *Helv. Chim. Acta.* **1991**, *35*, 397-406.

135. Worthington, V. L.; Fraser, W.; Schwalbe, C. H. *Carbohydrate Res.* **1995**, *275*, 275-284.

136. Hoffer, M. *Chem. Ber.* **1960**, *93*, 2777-2781.

137. Ogilvie, K. K.; McGee, D. P. C.; Biovert, S. M.; Hakimtahi, G. H.; Proba, Z. A. *Can. J. Chem.* **1983**, *61*, 1204-1213.

138. User's Manual for Applied Biosystems Model 392 DNA/RNA Synthesiser (Revision C, 1991).

139. Oliver, R. W. A. *HPLC of Macromolecules-A pratical Approach* **1989**, IRL Press: Oxford.

140. Haúpt, W.; Pingoud, A.; *J. Chromatography* **1983**, *260*, 419-427.

141. Fritz, H. -J.; Belagaje, R.; Brown, L.E.; Fritz, R. H.; Jone, R. A.; Lee, R. G.; Khorana, H. G. *Biochemistry* **1978**, *17*, 1257-1267.

142. Zhu, L.; Parr, G. R.; Fitzgerald, M. C.; Nelson, C. M.; Lloyd, M. S. *J. Am. Chem. Soc.* **1995**, *117*, 6048-6056

143. AlPrisco, D. D.; Marky, L. A.; Breslauer, K. J.; Turner, D. H. *Biochemistry* **1981**, *20*, 1309-1413.

144. *Thermodynamic Data for Biochemistry and Biotechnology*; Breslauer, J. K. Ed.; Springer-Verlag: Berlin, 1986.

145. Lin, T. -S.; Guo, J. -Y.; Schinazi, R. F.; Chu, C. K.; Xiang, J. -N.; Prusoff, W. H. *J. Med. Chem.* **1988**, *31*, 336-340.

146. Lin, T. -S.; Chen, M. S.; McLaren, C.; Gao, Y. S.; Chazzouli, I.; Prusoff, W. H. *J. Med. Chem.* **1987**, *30*, 440-444.

147. Lin, T. -S.; Mancini, W. R. *J. Med. Chem.* **1983**, *26*, 544-548.

148. Lin, T. -S.; Gao, Y. S.; Mancini, W. R. *J. Med. Chem.* **1983**, *26*, 1691-1696.

149. Lin, T. -S.; Yang, J. -H.; Liu, M. -C.; Shen, Z. -Z.; Cheng, Y. -C.; Prusoff, W. H.; Brinbaum, G. I.; Giziewicz, J.; Ghazzouli, I.; Brakovan, V.; Feng, J. -S.; Hsiung, G. -D. *J. Med. Chem.* **1991**, *34*, 693-701.

150. Wallis, M. P.; Spiers, I. D.; Schwalbe, C. H.; Fraser, W. *Tetrahedron Lett.* **1995**, *36*, 3659-3762.

151. Kamaike, K.; Takahashi, M.; Utsugi, K.; Tomizuka, K.; Ishido, Y. *Tetrahedron Lett.* **1995**, *36*, 91-94.

152. Westman, E.; Stromberg, R. *Nucleic Acids Res.* **1994**, *22*, 2430-2431.

153. Wilson, C. C.; Low, J. N.; Tollin, P. *Acta Cryst.* **1984**, *C40*, 1712-1715,

154. Cristalli, G.; Vittori, S.; Eleuteri, A.; Volpini, R.; Camaioni, E.; Lupidi, G.; Mahmood, N.; Bevilacqua, F.; Palu, G *J. Med. Chem.* **1995**, *38*, 4019-4025.

155. Dick, A.; Wallis, M. P.; Tisdale, M. J.; Fraser, W. unpublished data.

156. Birnbaum, G. I.; Giziewicz, *J. Can. J. Chem.* **1987**, *65*, 2135-2139.

157. Zamecnik, P. C.; Goodchild, J.; Taguchi, Y.; Sarin, P. S. *Proc. Nalt. Acad. Sci. USA.* **1986**, *83*, 4143-4146.

158. Eckstein, F. *Angew. Chem.* **1983**, *6*, 431-439.

159. Eckstein, F. *Annu. Rev. Biochem.* **1985**, *54*, 367-402.

160. *Nucleic Acids: The Vectors of Life*; Miller, P. S.; Agris, C. H.; Blake, K. R.; Murakami, A.; Spitz, S. A.; Reddy, M.P.; Ts'O, P. O. P.; PJER Publishing Co: Dordrecht, 1983.

161. Ts'O, P. O. P.; Miller, P. S.; Aurelian, L.; Murakami, A.; Agris, C.; Blake, K. R.; Lin, S. -B.; Smith, C. C. *Annu. NY Acad. Sci.* **1988**, *507*, 220-224.

162. Iyengar, R.; Frey, P. A. *J. Am. Chem. Soc.* **1984**, *106*, 8309-8314.

163. Frey, P. A.; Sammons, R. D. *Science* **1985**, *228*, 541-545.

164. Jaroszewski, J. W.; Coben, J. S. *Adv. Drug Delivery Rev.* **1991**, *6*, 235-250.

165. Zamecnik, P. C.; Stephenson, M. L. *Proc. Natl. Acad. Sci. USA.* **1978**, *75*, 280-284.

166. Stephenson, M. L.: Zamecnik, P. C. *Proc. Natl. Acad. Sci. USA.* **1978**, *75*, 285-288.

167. Loke, S. L.; Stein, C.; Zhang, X. H.; Mori, K.; Nakanishi, M.; Subasinghe, C.; Cohern, J. S.; Neckers, L. M. *Proc. Natl. Acad. Sci. USA.* **1989**, *86*, 3474-3478.

168. Yakubov, L. A.; Deeva, E. A.; Zarytova, V. F.; Ivanova, E. M.; Ryte, A. S.; Yurchenko, L. V.; Vlassov, V. V. *Proc. Natl. Acad. Sci. USA.* **1989**, *86*, 6454-6458.

169. Stein, C. A.; John, L.; Tonkinson, J.; Zhang, L, -M, Yakubov, L.; Gervasoni, J.; Taub, R.; Roterberg, S. A. *Biochemistry* **1993**, *32*, 4855-4861.

170. Shoji, Y.; Akhtar, S.; Periasamy, A.; Herman, B.; Juliano, R.L. *Nucleic Acids Res.* **1991**, *20*, 5543-5550.

171. Zhao, Q.; Matson, S.; Herrera, C. J.; Fisher, E.; Yu, H.; Krig, A. M. *Antisense Research and Development* **1993**, *3*, 55-66.

172. Akhtar, S.; Juliano, R. L. *Proc. Am. Assoc. Cancer Res.* **1991**, *32*, 333-337.

173. *Biomembranes: Molecular Structure and Function*; Gennis, R. B. Ed.; Springer-Verlag: New York, 1989.

174. White, J.: Kielian, M.; Helenius, A. *Quart. Rev. Biophys.* **1983**, *15*, 151-195.

175. Farahbahksh, Z. T.; Baldwin, R. L.; Wisieski, B. J. *J. Biol. Chem.* **1987**, *262*, 2256-2261.

176. Zamecnik, P.; Stephenson, M. *Proc. Natl. Acad. Sci. USA.* **1978**, *75*, 280-284.

177. Cerruzzi, M; Draper, K.; Schwartz, J. *Nucleosides and Nucleotides* **1990**, *9*, 679-695.

178. Chin, D. J.; Green, G. A.; Zon, G.; Szoka, F. C.; Straubinger, R. M. *New Biologist* **1990**, *9*, 1091-1100.

179. Leonetti, J. P.; Mechti, N.; Degols, G.; Gagnor, C.; Lebleu, B. *Proc. Natl. Acad. Sci. USA.* **1991**, *88*, 2702-2706.

180. Holt, J. T.; Redner, R. L.; Nienhus, A. W. *Molec. Cell. Biol.* **1988**, *8*, 963-973.

181. Goodchild, J. In *Conference on Antisense RNA and DNA*, St. Johns College, Mar 29-30, **1989**, Cambridge, U.K.

182. Harel-Bellan, A.; Ferris, D. K.; Vinocour, M.; Holt, J. T.; Farrar, W. L. *J. Immunol.* **1988**, *140*, 2431-2435.

183. Wickstorm, E. J.; *J. Biochem. Biophys. Methods* **1986**, **13**, 97-102.

184. Gao, W. -Y.; Han, F. S.; Storm, C.; Egan, W.; Cheng, Y. -C. *Mol. Pharmacol.* **1992**, *41*, 223-229.

185. Maher, L. J., III; Dolnick, B. J. *Nucleic Acids Res.* **1988**, *16*, 3341-3358.

186. Maksukura, M.; Shinozuka, K.; Zon, G.; Mitsuya, H.; Reitz, M.; Cohen, J. S.; Broder, S. *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 7706-7711.

187. Frey, P. A.; Sammons, R. D. *Science* **1985**, *228*, 541-545.

188. Iyengar, R.; Eckstein, F.; Frey, P. A. *J. Chem. Soc.* **1984**, *106*, 8309-8310.

189. Laplanche, L. A.; James, T. L.; Powell, C.; Wilson, W. D.; Uznanski, B.; Stec, W. J.; Summers, M. F.; Zon, G. *Nucleic Acids Res.* **1986**, *14*, 9081-9093.

190. Bower, M.; Summers, M. F.; Powell, F. C.; Shinozuka, K.; Regan, J. B.; Zon, G.; Wilson, W. D. *Nucleic Acids Res.* **1987**, *15*, 4915-4930.

191. Iyer, R.P.; Yu, D.; Agrawal, S. *Bioorg. Med. Chem. Lett.* **1994**, *4*, 2471-2476.

192. Greenberg, M. M. *Tetrahedron* **1995**, *51*, 29-38.

193. Greenberg, M. M.; Gilmore, J. L. *J. Org. Chem.* **1994**, *59*, 746-753.

194. Dreef-Tromp, C. M.; Hoogerhout, P.; van der Marel, G. A.; Van Boom, J. H. *Tetrahedron Lett.* **1990**, *31*, 427-430.

195. Dreef-Tromp, C. M.; van Dam, E. M. A.; van den Elst, H.; Van der Marel, G. A.; van Boom, J. H. *Nucleic Acids Res.* **1990**, *18*, 6491-6495.

196. Lyttle, M. H.; Wright, P. B.; Sinha, N. D.; Bain, J. D.; Chemberlain, A. R. *J. Org. Chem.* **1991**, **56**, 4608-4615.

197. Wada, T.; Sekine, M. *Tetrahedron Lett.* **1994**,*35*, 757-760.

198. Danishesky, S. J.; McClure, K. F.; Randolf, J. T.; Ruggeri, R. B. *Science* **1993**, *260*, 1307-1309.

199. Routledge, A.; Wallis, M. P.; Ross, K. C.; Fraser, W. *Bioorg. Med. Chem. Lett.* **1995**, *18*, 2059-2064.

200. Saha, A. K.; ßardara, M.; Waychunas, C.; Delecki, D.; Kutny, R.; Cavanaugh, P.; Yawman, A.; Upson, D. A.; Kruse, L. I. *J. Org. Chem.* **1993**, *58*, 7827-7831.

201. Cormier, J. F.; Ogilvie, K.K. *Nucleic Acids Res.* **1988**, *16*, 4583-4595.

202. Walsh, A. J.; Ross, K. C.; Routledge, A.; Wallis, M. P.; Fraser, W. *Pharmaceutical Sciences* **1996**, *2*, 33-38.