

University of Southampton Research Repository ePrints Soton

Copyright © and Moral Rights for this thesis are retained by the author and/or other copyright owners. A copy can be downloaded for personal non-commercial research or study, without prior permission or charge. This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the copyright holder/s. The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the copyright holders.

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given e.g.

AUTHOR (year of submission) "Full thesis title", University of Southampton, name of the University School or Department, PhD Thesis, pagination

UNIVERSITY OF SOUTHAMPTON

FACULTY OF ENGINEERING, SCIENCE AND
MATHEMATICS

School of Chemistry

**DNA Triplexes in Chemistry,
Biology and Medicine**

by

Radha Roxanna Tailor

Thesis for the degree of Doctor of Philosophy

September 2010

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF ENGINEERING, SCIENCE AND MATHEMATICS

SCHOOL OF CHEMISTRY

DOCTOR OF PHILOSOPHY

DNA TRIPLEXES IN CHEMISTRY, BIOLOGY AND MEDICINE

Radha Roxanna Tailor

The formation of DNA triple helices offers the possibility of selectively targeting specific genes to control their expression *in vivo*. This anti-gene strategy provides powerful tools for the development of therapeutics (anti-cancer drugs, drugs for viral infections) at the transcriptional level.

DNA triplexes are formed when an oligonucleotide binds to the major groove of double helical DNA; the third strand can bind in either a parallel motif, or an anti-parallel motif. The requirement of low pH for the protonation of cytosine in the parallel binding motif makes the formation of triple helices difficult under physiological conditions.

Described in this thesis is a novel method for the synthesis of the deoxycytidine analogue, 2-amino-3-methyl-5-(2'-deoxy-β-D-ribofuranosyl)pyridine (^{Me}P). The phosphoramidite monomer of ^{Me}P was synthesised and incorporated as a “protonated” cytidine analogue into triplex forming oligonucleotides (TFOs). It was compared with other cytosine analogues, 5-methyl-(2'-deoxy-β-D-ribofuranosyl)cytosine (^{Me}C), 2'-O-methyl ^{Me}P (^{Me}P_{OMe}), and 2'-O-aminoethyl ^{Me}C (^{Me}C_{AE}). Triplex stability studies indicate that over the pH range 6.2-8.0, the general trend observed in terms of melting temperature (T_m) was as follows: ^{Me}P > ^{Me}C > ^{Me}P_{OMe} > ^{Me}C_{AE}. DNase I footprinting studies indicate that at pH 7.5, ^{Me}P, when incorporated into the TFO, enhances the stability of the triplex by three-fold relative to ^{Me}C. In addition, UV melting, DNase I footprinting, and gel electrophoresis studies were carried out on a triplex formed by the binding of a TFO containing ^{Me}P and a 5'-Psoralen to a target duplex. This revealed the benefits of the combined modifications on the stability of the resultant triplex. “Soaking” experiments (*in vivo*) were also performed with this TFO on the organism *C. elegans* (the worms were soaked in solutions of the TFO for TFO delivery), to observe whether the TFO would induce loss-of-function phenotypes. T_m measurements indicated that in the pH range 6.6-8.0, photo-crosslinking of the TFO to the duplex created a shift in the triplex T_m of ~ + 26 °C when compared to the un-crosslinked triplex.

CONTENTS

Abstract	i
Contents	iii
Declaration	vii
Acknowledgments	ix
Abbreviations	xi
1. Introduction	1
1.1 Genomes	3
1.2 The History of DNA Discovery	3
1.3 DNA and RNA Structures	4
1.3.1 Primary Structure	4
1.3.2 Secondary Structure	10
1.4 DNA Function and the Genetic Code	15
1.4.1 DNA Replication	16
1.4.2 DNA Transcription	17
1.4.3 RNA Translation	18
1.4.4 The Genetic Code and the Wobble Hypothesis	19
1.4.5 DNA Mutations and Repair	20
1.4.6 Reverse Transcription	23
1.5 The Antisense Approach to Reverse Transcription and Gene Inhibition	24
1.6 DNA Triple Helices and the Antigene Approach to Gene Inhibition	26
1.6.1 The Discovery of and an Introduction to DNA Triple Helices	26
1.6.2 Triplex Motifs and their Structure	30
1.6.3 Limitations to Triple Helix Formation	33
1.6.4 Methods to Combat the Limitations to Triple Helix Formation	36
1.7 Methods to Study the Biophysical Properties of Nucleic Acids	72
1.7.1 Ultra-Violet (UV) Melting	72
1.7.2 Footprinting	76

1.8	Aims of this PhD Research	78
1.9	Psoralen	79
1.10	Psoriasis	81
1.11	Caenorhabditis Elegans (<i>C. elegans</i>)	82
2.	Synthesis of the C-nucleoside Cytidine Analogue 3-methyl-2-aminopyridine deoxyriboside (^{Me}P)	85
2.1	Introduction	87
2.2	C-nucleosides	89
2.2.1	Introduction to C-nucleosides	90
2.2.2	Synthesis of C-nucleosides	90
2.3	Synthesis of 3-methyl-2-aminopyridine (^{Me}P)	105
2.3.1	Background Work	105
2.3.2	The Approaches Towards ^{Me} P	110
3.	Biophysical and Biological Studies	141
3.1	Initial Studies on ^{Me}P	143
3.1.1	<i>In Vitro</i> studies	143
3.2	Comparison of Cytosine-Nucleoside Analogues in TFOs	148
3.2.1	<i>In Vitro</i> Studies	148
3.2.2	Conclusion	167
3.3	Psoralen Photo-Crosslinking Reactions	168
3.3.1	<i>In Vitro</i> Studies	168
3.3.2	<i>In Vivo</i> Studies	178
4.	Conclusions	181
4.1	The Achievements of this Research	183
4.2	Research Continued	186
4.3	Applications for Gene Therapy?	187
4.4	Biological Barriers to Gene Therapy	188
4.5	Moving Forward	188
5.	Experimental	191
5.1	Synthesis	193
5.1.1	General Experimental	193
5.2	Compounds Synthesised	195
5.3	Oligonucleotide Synthesis	279

5.3.1	Synthesis of Oligonucleotides	279
5.3.2	Oligonucleotide Purification	280
5.3.3	Oligonucleotide Analysis	280
5.3.4	Oligonucleotides Synthesised	281
5.4	Biophysical Studies	284
5.4.1	Ultraviolet-Melting Analysis	284
5.4.2	Footprinting Experiments	287
6.	References	289
7.	Appendix	305
7.1	Publications	307

Declaration of Authorship

I, Radha Roxanna Tailor declare that the thesis entitled DNA Triplexes in Chemistry, Biology and Medicine and the work presented in the thesis are both my own, and have been generated by me as the result of my own original research. I confirm that:

- this work was done wholly or mainly while in candidature for a research degree at this University;
- where any part of this thesis has previously been submitted for a degree or any other qualification at this university or any other institution, this has been clearly stated;
- where I have consulted the published work of others, this is always clearly attributed;
- where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Radha. R. Tailor, September 2010

Acknowledgements

I would like to thank my supervisor Prof. Tom Brown for allowing me this great opportunity to work on a very interesting PhD project. I would also like to thank him for his endless support, ideas and valuable advice he has provided me throughout my postgraduate studies.

Secondly, I would like to express my gratitude to past and present members of the Brown research group that have created a pleasurable working environment, and for the scope of fresh ideas for my research. I would also like to thank those who proof read this thesis.

I would like to thank Prof. Keith Fox and Dr. Nuria Vergara for their input with the biophysical analysis involved within this PhD.

Thanks to ATDBio for oligonucleotide synthesis and purification, Dr John Langley and Miss Julie Herniman for access to the mass spectrometry facilities, and Dr Neil Wells and Mrs Joan Street for the use of the NMR systems.

Finally, thank you to my family (Bhasker, Kiran, Krishna, and Chrishan Tailor) and Earl Peters, for their unconditional love, support, motivation and encouragement throughout my PhD. Thank you dad for proof-reading my thesis so thoroughly and giving your valuable suggestions.

Abbreviations

A	Adenine
A	Absorbance in the Beer-Lambert Law
Ab _S Max	Absorbance maximum
α	Alpha
Aq	Aqueous
β	Beta
BAU	Bis amino dU
BBr ₃	Boron tribromide
BCl ₃	Boron trichloride
BNA	Bridged nucleic acid
BuLi	Butyl lithium
C	Cytosine
^{Me} C	5-methyl 2'-deoxy Cytosine
^{Me} C _{AE}	5-methyl-(2'-O-aminoethyl- β -D-ribofuranosyl)cytosine
CHCl ₃	Chloroform
CDCl ₃	Deuterated chloroform
CD ₃ CN	Deuterated Acetonitrile
CH ₃ CN	Acetonitrile
COSY	Correlation spectroscopy
δ	Chemical shift in parts per million
dA	2'-deoxyadenosine
dC	2'-deoxycytidine
DCM	Dichloromethane
DEAD	Diethylazodicarboxylate
DEED	<i>N,N</i> -diethyl-ethylenediamine
DEPT	Distortionless Enhancement through Polarization Transfer
dG	2'-deoxyguanosine
DIPEA	<i>N,N</i> -Diisopropylethylamine
DMAP	<i>N,N</i> -dimethylaminopyridine
DMF	<i>N,N</i> -dimethylformamide

DMSO	Deuterated dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dsDNA	Double stranded DNA
DNG	Deoxyribonucleic guanidine
dNTP	Deoxyribonucleotide triphosphate
D ₂ O	Deuterated water
EA	Ethyl acetate
EDTA	Ethylenediaminetetraacetic acid
eq	Equivalent
ES	Electrospray
Et ₂ O	Diethyl ether
EtOH	Ethanol
Et ₃ N	Triethylamine
G	Guanine
H	hexa(ethylene glycol)
HBr	Hydrobromic acid
HCl	Hydrochloric acid
Hex	Hexane
H ₂ O	Water
HMQC	¹ H- ¹³ C correlation NMR spectroscopy
HPLC	High pressure liquid chromatography
HRMS	High resolution mass spectrometry
Hz	Hertz
ΔH	The change in enthalpy
ⁱ PrOH	isopropanol
IR	Infrared spectroscopy
<i>J</i>	Coupling constant
KCl	potassium chloride
<i>K_d</i>	Rate constant of dissociation
<i>K_a</i>	Rate constant of association
μl	Microlitre
λ _{Max}	Lambda maximum
LNA	Locked nucleic acid

LRMS	Low resolution mass spectrometry
MALDI-TOF	Matrix assisted laser desorption ionisation-time of flight
MeCN	Acetonitrile
MeOH	Methanol
MeNH ₂	Methylamine
mer	Oligonucleotide length in terms of the number of nucleotide units
MHz	Megahertz
m. p.	Melting point
mRNA	Messenger RNA
MS	Mass spectrometry
M/Z	Mass/charge
NAP	Nucleic acid purification (a disposable sephadex column)
NMR	Nuclear magnetic resonance
Oligo	Oligonucleotide
Me^eP	3-methyl-2-amino-5-(2'-deoxy-β-D-ribofuranosyl)pyridine
Me^eP_{OMe}	3-methyl-2-amino-5-(2'-O-methyl-β-D-ribofuanosyl)pyridine
Pac	Phenoxyacetyl
pH	Potential of hydrogen (negative logarithm of the positive hydrogen ion concentration)
Ph ₃ P	Triphenylphosphine
p <i>K_a</i>	Ionisation constant (negative logarithm of the acid dissociation constant, <i>K_a</i>)
PMB	paramethoxybenzyl
PNHDMAP	dimethylaminopropyl phosphoramidate linkages
Pso	Psoralen
pu	Purine
py	Pyrimidine
Pyr- <i>d</i> 5	Deuterated pyridine
R _f	Retention factor
RNA	Ribonucleic acid
RNase-H	Ribonuclease-H
rRNA	Ribosomal RNA
RNG	Ribonucleic guanine

r.t	Room temperature
S_{ME}	2'- <i>O</i> -methoxyethyl S
ssDNA	Single stranded DNA
ssRNA	Single stranded RNA
T	Thymidine
T_{AE}	2'- <i>O</i> -aminoethyl thymidine
TBAF	Tetrabutyl ammonium fluoride
TFA	Trifluoroacetic acid
TFAA	Trifluoroacetic anhydride
TFO	Triplex-forming oligonucleotide
THF	Tetrahydrofuran
TiPDS	1,1,3,3-Tetraisopropylidisiloxane
TLC	Thin layer chromatography
T _m	Temperature of melting
TMS	Trimethylsilyl
Tol	Toluy group
Tris-HCl	Tris(hydroxymethyl)aminomethane hydrochloride
tRNA	Transfer RNA
U	Uracil
UV	Ultraviolet radiation of the electromagnetic spectrum
% wt/vol	Percentage weight to volume

CHAPTER 1

Introduction

1. INTRODUCTION

1.1 GENOMES¹

In all organisms, the biological information that is needed to construct and maintain that organism is stored within the genome. Most genomes (human genome and those of other cellular life forms), are made up of DNA (deoxyribonucleic acid), however, a few viruses have RNA (ribonucleic acid) genomes. Nucleotides are the monomeric subunits that make up the long thread-like polymeric chains of DNA and RNA molecules.

The human genome consists of two distinct parts:

(i) The **nuclear genome/nuclear DNA** – located in the nucleus of a cell, and is made up of approximately 3,200,000,000 nucleotides of DNA, which are distributed amongst 23 pairs of chromosomes; These 23 pairs of chromosomes consist of 22 autosomal pairs and the sex chromosome pair, XX for females and XY for males. In each pair, one chromosome is paternally inherited, the other maternally inherited.

(ii) The **mitochondrial genome/mitochondrial DNA (mtDNA)** - is a circular DNA molecule of 16,569 nucleotides, multiple copies of which are located in the mitochondria. They are located within the cell, surrounding the nucleus. In humans, 100-10,000 separate copies of mitochondrial DNA are usually present per cell. In most multicellular organisms, mtDNA is maternally inherited (inherited from the mother).

Each of the cells (approximately 10^{13}) in the adult human body has its own copy or copies of the genome, the only exceptions being those few cell types that lack a nucleus in their fully differentiated state (red blood cells).

1.2 THE HISTORY OF DNA DISCOVERY^{1,2}

It has been more than 200 years since DNA was first discovered. During this time, many dedicated scientists and geneticists have contributed to the current knowledge on the human genome.

In 1866, Gregor Mendel established the basic laws of inheritance. He discovered that genes govern hereditary traits. They exist in pairs, and undergo separation, independent assortment, and persist unchanged through successive generations (hereditary transmission).¹

DNA was first discovered in 1869 by Johann Friedrich Miescher.¹ His initial research focused on determining the fundamentals governing human leukocyte cells (white blood cells) which were obtained from human pus. Miescher had obtained the first comparatively clean preparation of DNA/‘nuclein’ (as he had originally named it). Using elemental analysis, he detected various elements typically found in organic molecules: carbon, hydrogen, oxygen, nitrogen. Meischer’s analyses also revealed that nuclein contained a large proportion of phosphorus.²

In 1889, Richard Altmann succeeded in the complete separation of DNA from proteins. As a result of its acidic behaviour, he named the substance ‘nucleic acid’. ‘Nucleic acid’ was in fact the same substance Meischer had isolated and first described as ‘nuclein’.²

Albrecht Kossel confirmed that nuclein/DNA is restricted to the nucleus. He also confirmed that the purine and pyrimidine bases, one pentose sugar, and phosphoric acid, are the fundamental building blocks. He also deduced from his research that nucleic acids are involved in the synthesis of new protoplasm during growth and replacement. In 1910 Kossel was awarded the Nobel Prize in Physiology or Medicine, in recognition of his groundbreaking work on the chemistry of proteins and nucleic acids. Despite these early breakthroughs, the full importance and impact of nucleic acids remained unclear for several decades.²

1.3 DNA AND RNA STRUCTURE³

1.3.1 Primary Structure³

Nucleic acids are made up of a linear array of nucleotides. ‘Nucleotides’, a term that was introduced by Levene⁴ and Jacobs in 1909, are the phosphate esters of nucleosides. Both RNA (ribonucleic acid) and DNA (deoxyribonucleic acid) are composed of

nucleotides. Whilst conducting hydrolysis studies of nucleic acids they discovered that the carbohydrate component of DNA was 2'-deoxyribose, and that of RNA was D-ribose.^{5,6} DNA therefore contains 2'-deoxyribonucleotides and RNA contains ribonucleotides.³

As determined by Kossel, all nucleotides consist of three major components: (1) a nitrogen heterocyclic base; (2) a pentose sugar; and (3) a phosphate residue. The major bases are monocyclic pyrimidines or bicyclic purines. Adenine (A) and guanine (G) are the major purine derivatives, and are found to be present in both DNA and RNA. Cytosine (C), Thymine (T) and uracil (U) are the major pyrimidine derivatives. Both C and T are present in DNA while in RNA, T is replaced by its non-methylated pyrimidine derivative, U (*fig. 1.1*).³

Nucleosides are the condensation products of a pentose sugar and a nitrogenous base. In RNA, the nucleobase is joined from either the N-1 of the pyrimidine or the N-9 of the purine, to the C-1' position of the D-ribose sugar (*fig. 1.2*). This bond locks the D-ribose (which can mutarotate under certain conditions, adopting furanose, acyclic or pyranose forms; *fig. 1.3*) into a five-membered furanose ring and resides on the same side of the sugar ring as the C-5' hydroxymethyl, forming a β -glycosidic linkage. In DNA, the condensation reaction occurs between the same nucleobase nitrogen as in RNA, however, with the C-1' position of 2'-deoxy-D-ribose (*fig. 1.4*).³

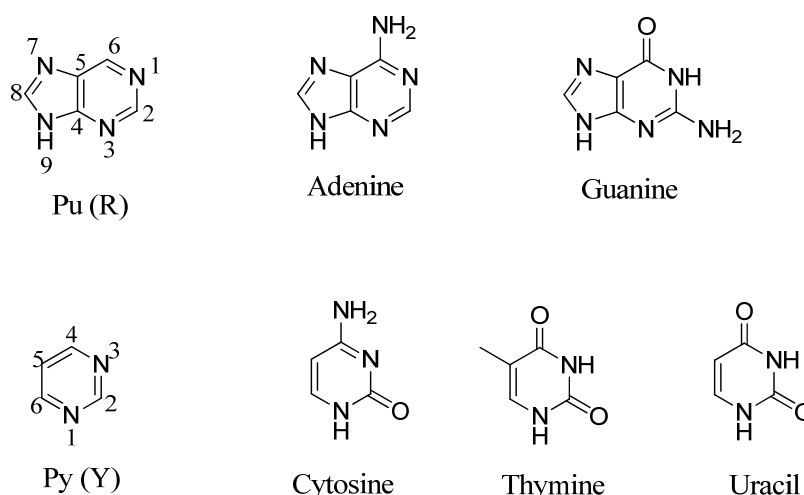


Fig. 1.1: Structures of the five major purine and pyrimidine bases of nucleic acids in their dominant tautomeric forms.

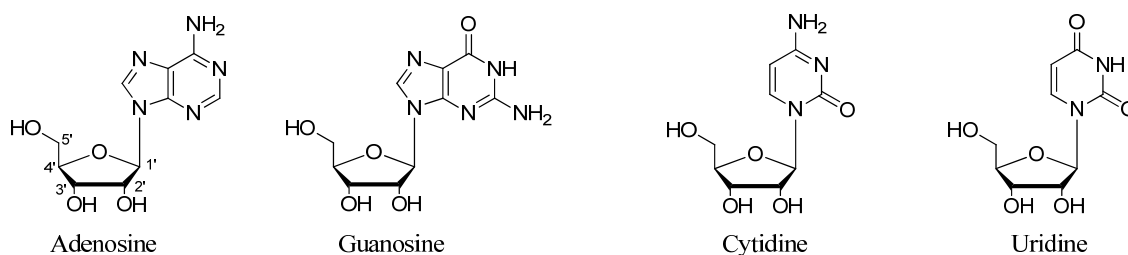


Fig. 1.2: Structures of the four ribonucleosides.

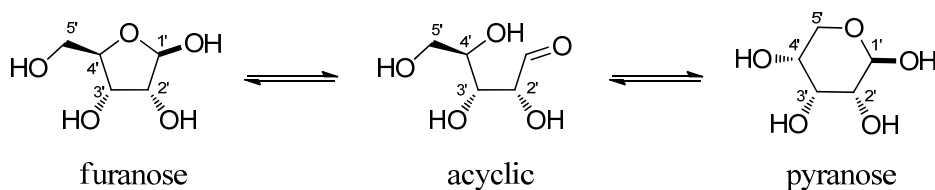


Fig. 1.3: Mutarotation structures of D-Ribose.

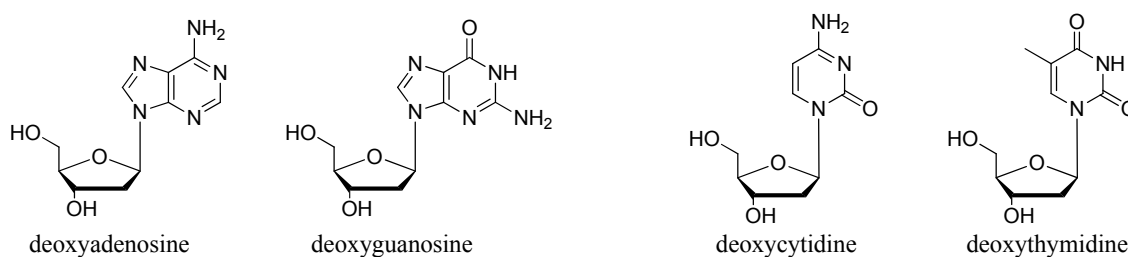


Fig. 1.4: Structures of the four deoxyribonucleosides.

Rotation about the glycosidic bond allows the nucleobase to occupy either of two principal orientations, giving rise to two conformers. The nucleoside is referred to as the *anti* conformer when the smaller H-6 atom in pyrimidines, or H-8 atom in purines resides above the sugar ring. Whereas, when the larger O-2 atom from pyrimidines or the N-3 atom from purines resides above the sugar ring, the nucleoside is referred to as the *syn* conformer (*fig. 1.5*). In DNA, to enable base pairing, the *anti* conformation is required. It is also generally favoured on steric grounds.³

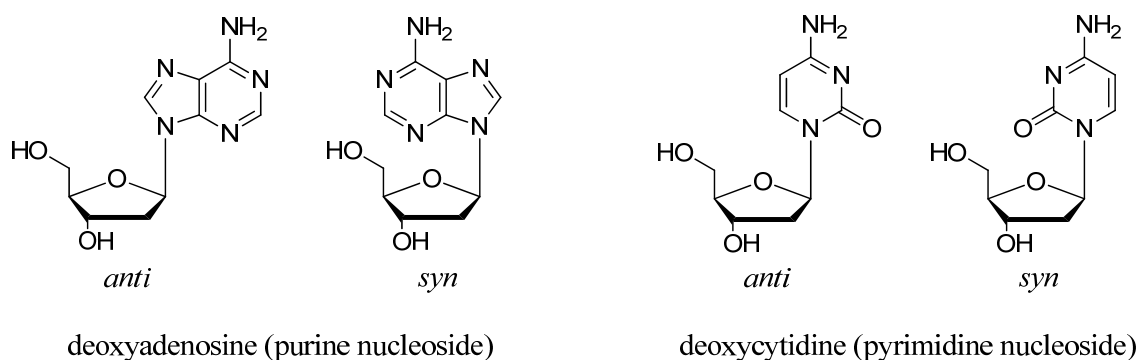


Fig. 1.5: *Anti* and *syn* conformers for a purine nucleoside (A) and a pyrimidine nucleoside (C).

The 1'-position of the furanose ring is an anomeric centre. When the nucleobase resides on the same face of the sugar ring as the 5'-hydroxyl group it is known as the β -anomer, when it resides on the opposite face it is known as the α -anomer (*fig. 1.6*). In both DNA and RNA, it was deduced from respective x-ray crystallography studies that the nucleosides exist in their β -configuration.³

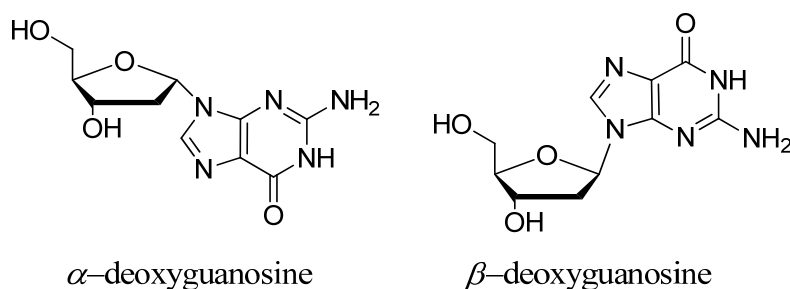


Fig. 1.6: The structures of α - and β -deoxyguanosine.

The furanose rings encounter a phenomenon described as 'puckering'. Puckering is the twisting of the sugar ring out of plane to minimise non-bonded steric and electronic interactions between the substituents. It identifies the displacement of carbons 2' and 3' from the plane $C^{1'}-O^{4'}-C^{4'}$. The sugar pucker is described as $C^{2'}$ -*endo* (south, S) when the *endo* displacement of $C^{2'}$ is greater than the *exo* displacement of $C^{3'}$ (*fig. 1.7*). Whereas the sugar pucker is described as $C^{3'}$ -*endo* (north, N) when the *endo* displacement of $C^{3'}$ is greater than the *exo* displacement of $C^{2'}$. The *endo* face of the furanose is on the same side as $C^{5'}$ and the base; the *exo* is on the opposite face to the base.^{7,8} In solution the N and S conformations are in dynamic equilibrium. There exists a number of influential factors that give preference for a specific conformation: (1) the

preference of electronegative substituents at C^{2'} and C^{3'} for axial orientation; (2) the orientation of the base (*syn* goes with C^{2'}-*endo*); (3) the formation of an intra-strand hydrogen bond from exocyclic OH^{2'} in one RNA residue to endocyclic O^{4'} in the next, which favours C^{3'}-*endo* pucker. DNA adopts an S-type conformation, while RNA adopts an N-type conformation.³

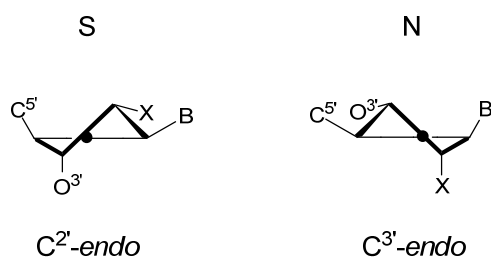


Fig. 1.7: The structures of the sugar pucker in β -nucleosides; C^{2'}-*endo* (S) and C^{3'}-*endo* (N). Where B = nucleobase; X = OH (RNA, N-type) or H (DNA, S-type).

Within a nucleotide, the phosphate group can be found at either the 3'- or 5'-position of the sugar, depending on the method used to break down the DNA to produce the nucleotide (*fig. 1.8*). It was P. A. Levene⁹ that also discovered during hydrolysis studies on nucleic acids, the order in which the three major components of a single nucleotide are presented (phosphate-sugar-base) in both DNA and RNA.³

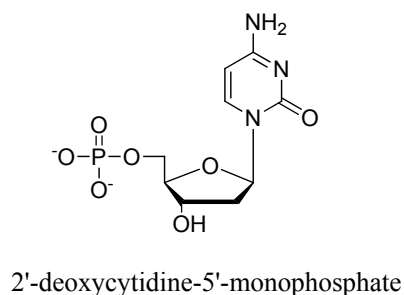


Fig. 1.8: The structure of the C nucleotide, 2'-deoxycytidine-5'-monophosphate.

DNA and RNA are linear polynucleotides in which each nucleoside is linked to the next by means of a 3'-5'-phosphodiester (*fig. 1.9*). Klein and Thannhauser established that each DNA nucleoside is joined by a phosphodiester linkage, from its 5'-hydroxyl group to the 3'-hydroxyl group of one neighbour, and by a second phosphodiester linkage from its 3'-hydroxyl group to the 5' hydroxyl of its other neighbour. The chain

possesses directionality (the order of a nucleic acid sequence is always written in the 5' to 3' direction). These phosphodiester residues are stable towards chemical hydrolysis, and are deprotonated at physiological pH resulting in nucleic acid chains being highly negatively charged, and consequently forming salts.³ The primary structure of DNA is given by the sequence of bases within a single strand. It is this sequence that gives DNA its uniqueness within individuals.

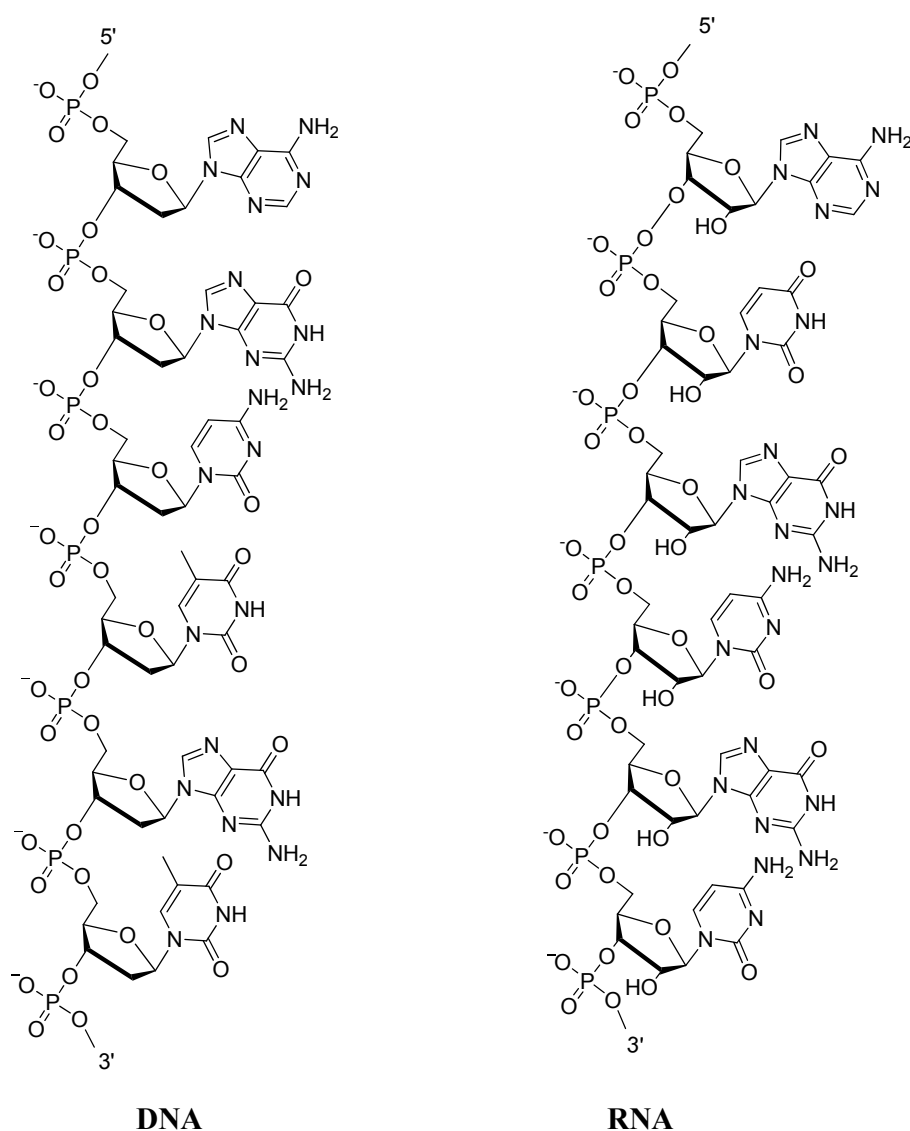


Fig. 1.9: The structures of single stranded DNA and RNA.

1.3.2 Secondary Structure

It was the combination of ideas and discoveries by Erwin Chargaff¹⁰⁻¹³, Rosalind Franklin^{14,15}, Maurice Wilkins, Linus Pauling, Sven Furberg¹⁶, James Watson¹⁷, Francis Crick¹⁷, and many other scientists of that time that led to the breakthrough of the secondary structure of DNA.

Erwin Chargaff¹⁰⁻¹³ analysed the base composition of DNA in a number of organisms. He reported that DNA composition varies from one species to another. He developed a method to separate the individual products of hydrolysis of DNA by means of paper chromatography and by use of ultraviolet spectroscopy, to quantify their relative abundance. He noticed regularity in the ratios of the bases. The ratio of the amounts of adenine to thymidine, and the ratio of guanine to cytosine, is always very close to unity. Hence the proportion of purines, (A + G), is always equal to the proportion of pyrimidines, (C + T). Although the ratio of (G + C)/(A + T) varies from species to species, different tissues from a single species give DNA of the same composition.

Sven Furberg¹⁶ reported in 1952 the possibility of the structure of nucleic acids adopting one of the two proposed models. In model I, 'the pyrimidine and purine rings are piled in a central column directly on top of each other, 3.4 Å apart, with the ribose rings and the phosphate groups in a spiral enclosing the column. The structure is held together by Van der Waals attraction between the flat rings in the column'. In model II, 'the ribose rings and phosphate groups form a flattish central column, from which the purines and the pyrimidines stand out perpendicularly. Successive bases are pointing outwards in opposite directions, all the N-C^{1'} bonds being roughly parallel, and there is one base directly above another at a distance of 6.8 Å. No forces are acting between the pyrimidine and purine bases, therefore, the whole molecule must be held together by interactions between atoms in the pentose rings and the phosphate groups, which are packed very efficiently in a compact sheet'. It was suggested that the available evidence from x-ray crystallography studies favoured model I.

Pauling and Corey^{18,19} using x-ray crystallographic studies and other principles of the molecular structure, proposed the structure of nucleic acids to be three intertwined

helical polynucleotide chains in a right-handed nature. The phosphate groups form the dense core and the bases project radially on the outside with their planes being approximately perpendicular to the molecular axis.

Rosalind Franklin and R. Gosling^{14,15}, using x-ray diffraction studies on DNA fibres, identified that the hydrophilic phosphate backbones of the nucleotide chains of DNA should be positioned so as to interact with water molecules on the outside of the molecule while the hydrophobic bases should be packed into the core. This conclusion was reached from the studies by Gulland *et al.*²⁰, they showed that in aqueous solution the CO and NH₂ groups of the bases are inaccessible and cannot be titrated, whereas the phosphate groups are fully accessible. Also, the ready availability of the phosphate groups for interaction with proteins can most easily be explained in this way.

James Watson and Francis Crick¹⁷ were the first scientists to formulate an accurate description of DNA. Their work was directly dependent on the research of numerous scientists before them that have been discussed above. The structure they suggested contained two helical chains each coiled round the same axis (*fig. 1.10*)²¹. Each chain consists of phosphodiester groups joining β -D-deoxyribofuranose residues with 3',5' linkages. Both chains follow right-handed helices, the sequence of residues in the two chains run in opposite directions; one strand runs 3' to 5' and the other 5' to 3'. Each chain loosely resembles Furberg's model I¹⁶ discussed above: the bases are on the inside of the helix and the phosphates on the outside. The sugar is roughly perpendicular to the attached base. There is a residue on each chain every 3.4 Å, in the z-direction (i.e. the bases stack one upon another). An angle of 36 ° exists between adjacent residues, hence the structure repeats after ~10 residues on each chain or after 34 Å vertically. The radius of the helix is 10 Å. The phosphates located on the outside are easily accessed by cations. DNA is an open structure with a high water content. The two chains are held together by hydrogen bonds that exist between the purine and pyrimidine bases; a single base from one chain is hydrogen-bonded to single base from the other chain, the two lie side-by-side with identical z-coordinates with their planes being perpendicular to the fibre axis. Hence, when the sequence of bases on one chain is given, the sequence on the opposite chain is automatically determined; the two strands are complementary. The stability of the duplex is derived from both base stacking (π - π interactions of the stacked aromatic bases) and hydrogen bonding.

The bases are present in their amino-keto rather than enol tautomeric forms. Adenine hydrogen bonds to thymidine, and guanine hydrogen bonds to cytosine (*fig. 1.11*); known as Watson-Crick base pairing and agrees with Erwin Chargaff's base abundance findings¹³. The NH groups of the bases are good hydrogen bond donors (**d**), while the sp^2 -hybridised electron pairs on the oxygens of the base C=O groups and on the ring nitrogens are good hydrogen bond acceptors (**a**). The **a**·**d** hydrogen bonds are largely electrostatic in character. There exist two hydrogen bonds in an A·T base pair and three in a G·C base pair. The formation of hydrogen bonds between the base pairs gives a C^{1'}-C^{1'} distance of $10.60 \pm 0.15 \text{ \AA}$, with an angle of $68 \pm 2^\circ$ between the two glycosidic bonds. As a result of this, the base pairs are pseudosymmetric, so when one base pair is overlaid by the other base pairs the phosphodiester backbones fall on top of each other, and hence all four base pairs fit neatly within the double helix.

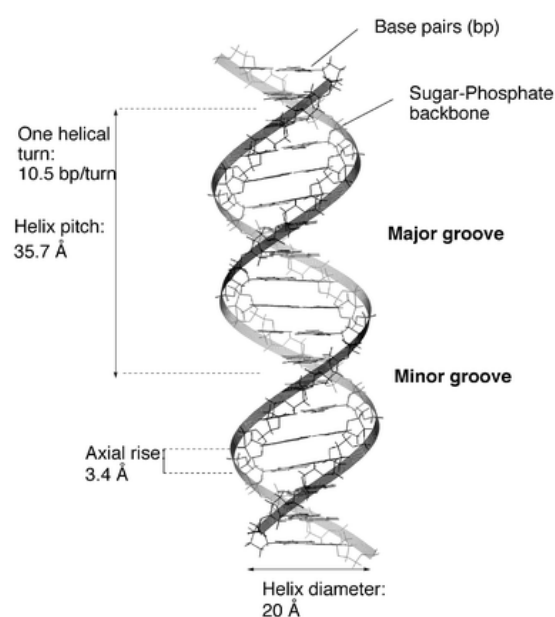


Fig. 1.10: The structure of the DNA double helix along with the parameters.²¹

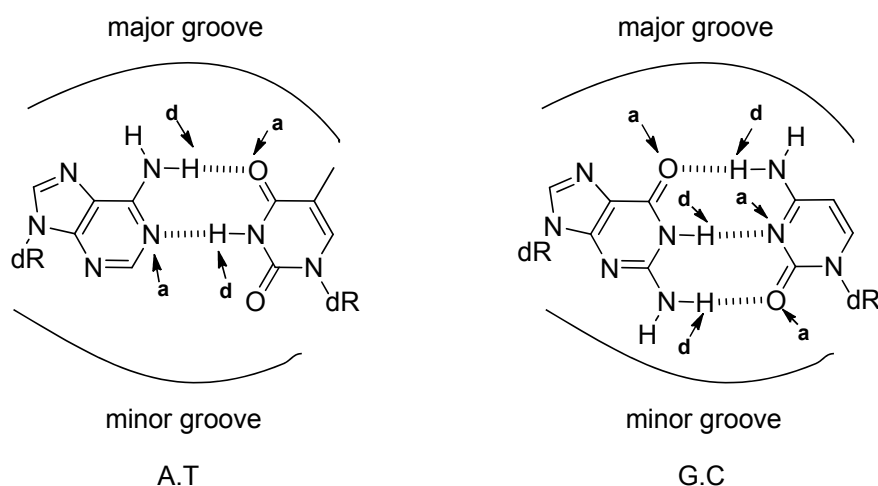


Fig. 1.11: Structures of the Watson-Crick base pairs; A·T and G·C respectively, indicated are the regions of the major and minor grooves along with the hydrogen bond donors (**d**) and acceptors (**a**). In each case dR indicates the deoxyribose sugar.

Watson-Crick base-pairing is the dominant kind of interaction between bases; however, other pairings, such as Hoogsteen pairs exist (*fig. 1.12*). Hoogsteen pairs have an angle of 80° between the glycosidic bonds and an 8.6 \AA separation of the anomeric carbons; hence they are not pseudosymmetric with Watson-Crick pairs. Reversed Hoogsteen pairs and reversed Watson-Crick pairs also exist, they occur when one base is rotated through 180° relative to the other.

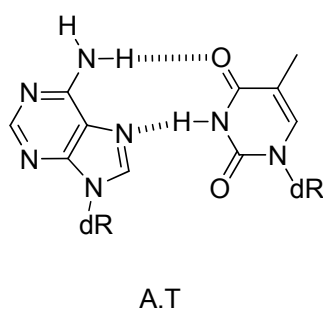


Fig. 1.12: Hoogsteen base pair A·T.

On further studies of heterogeneous DNA fibres, Franklin and Gosling^{14,15} identified 2 distinct conformations for the DNA double helix which can interconvert. The highly crystalline A-DNA (*fig. 1.13*)²¹ is the favoured form at low humidity (and high salt), while the less-ordered B-DNA, for which the properties were reported by Watson and Crick and are discussed above, is the dominant structure at high humidity (and low salt).

In the A-form DNA, two co-axial helical chains run in opposite directions in a right-handed manner, the number of residues per helical turn increases from ten (in B-DNA) to eleven, hence the pitch of the helix decreases from 34 Å to 28 Å. The vertical component of the inter-base distance is therefore decreased from 3.4 Å in B-DNA to 2.55 Å in A-DNA. The planes of the base pairs are displaced off-axis towards the minor groove and are tilted at $\sim 25^\circ$. In A-DNA the phosphate groups lie on a helix of radius 9 Å. The furanose ring has C^{3'}-*endo* pucker and the glycosidic bond is in the *anti* conformation. The major groove of A-DNA is extremely deep and the minor groove is extremely shallow. The sugars of B-DNA (the principal form) have C2'-*endo* pucker, with all the glycosides having *anti* conformation. B-DNA has a wide major groove and a narrow minor groove, of equal depth, running around the helix along the entire length of the molecule. Proteins interact with DNA principally in the major groove, and some small drug molecules (e.g. netropsin, distamycin) bind in the minor groove.

RNA can also form right-handed duplexes, which exist in the A-form. The same base pairing rules apply; A·U and G·C. There are a number of other conformations of RNA and DNA, most of which are subtle variations on the A- and B- form; they are classed according to their sugar pucker; C^{3'}-*endo* for the A-form, and C^{2'}-*endo* for the B-form.³

Z-DNA was discovered from x-ray crystal structures on the chemically synthesised strand d(CGCGCG) which spontaneously forms a duplex in aqueous buffer, and exists when the helix strands run antiparallel, but left-handed in nature. There are 11.6 base pairs per helical turn, the major groove is completely flattened out on the surface of the molecule and the minor groove is very deep. The base pairs are displaced off-axis. There exists an alternation of *syn* and *anti* conformations at guanine and cytosine respectively which produces a zig-zag chain path and hence a dinucleotide unit repeat.²¹

In DNA and RNA, the structure conformation and properties are governed by the sequence of base pairs within. It is the π - π stacking interactions between successive base-pairs that provide this crucial link between sequence, structure and properties. There are four interactions that contribute towards the energy of π - π interactions: (1) van der Waal's interactions; (2) electrostatic interactions between partial atom charges; (3) electrostatic interactions between the negatively charged π -electron clouds out of the

planes of the aromatic molecules with the positively charged σ -nuclei framework; and (4) electrostatic interactions between the π -electron clouds and partial atomic charges.²² Furthermore, the stability of the DNA or RNA structure is also governed by van der Waals' interactions, electrostatic interactions, solvent interactions, and hydrogen bonding between base pairs.²²

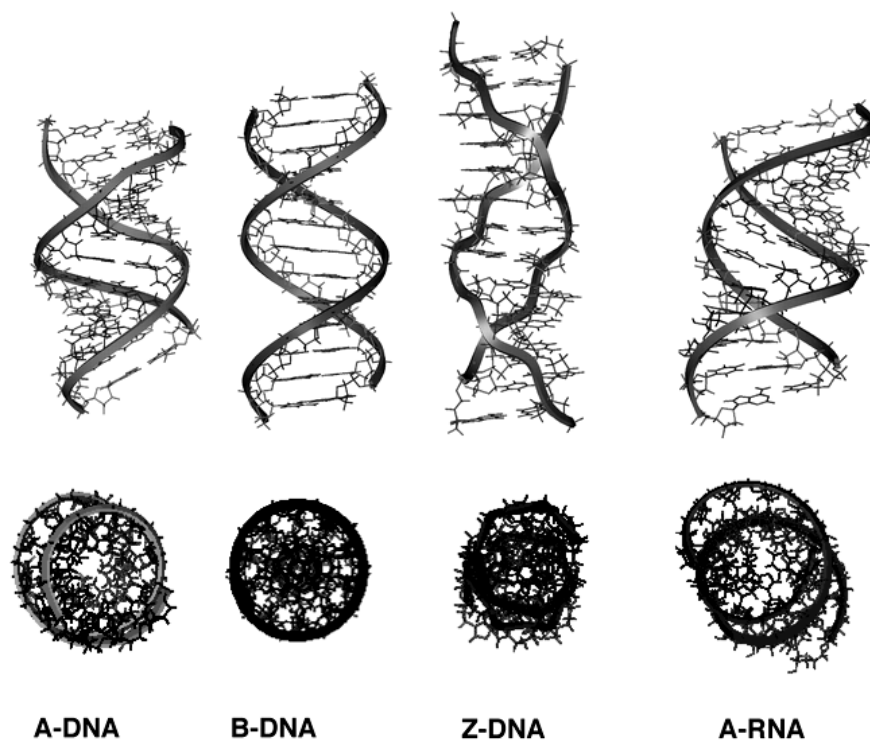


Fig. 1.13: The major nucleic acid duplex conformations in DNA (A-form, B-form, and Z-form) and RNA (A-form). The sugar-phosphate backbone is represented by the ribbon.²¹

1.4 DNA FUNCTION AND THE GENETIC CODE

Watson and Crick observed that the specific base-pairing within DNA could provide a mechanism for heredity. Genetic information is encoded in the sequence of bases that are attached to the C^{1'} of deoxyribose sugars along a DNA chain. This specific sequence determines the structure of proteins and represents genetic inheritance. Overall, DNA codes for RNA, which codes for proteins.

1.4.1 DNA Replication²³⁻²⁵

An exact copy of DNA is created prior to cell division; the enzyme topoisomerase unwinds the DNA at the origin, and helicase splits the strands of the DNA molecule apart by breaking the hydrogen bonds between base-pairs, forming a ‘bubble’ within the ‘mother’ duplex (*fig.1.14*). Origins tend to be where there are A.T-rich regions due to there existing two hydrogen bonds between the base pairs rather than three, which exist between G.C base pairs, hence a positive relationship between the number of hydrogen bonds and the difficulty in breaking these bonds. RNA primase creates RNA primer nucleotides at the initiation point on each template strand; the leading strand receives one RNA primer per active origin of replication while the lagging strand receives several. DNA polymerase catalyses the polymerisation of the deoxyribonucleotides into a complementary DNA strand at the 3'-end of the RNA primer, by adding new deoxyribonucleotide triphosphates (dNTP) complementary to the template strand by Watson-Crick pairing, one at a time *via* the creation of phosphodiester bonds. Extension of the leading strand occurs in one continuous motion along the template in a 3' to 5' manner, while the lagging strand is extended in a discontinuous motion due to Okazaki fragments²⁶ (relatively short fragments of DNA with no RNA primer at the 5'-end) in a 5' to 3' direction. Due to its orientation, opposite to the working orientation of DNA polymerase, which moves on a template in a 3' to 5' manner, replication of the lagging strand is more complicated than that of the leading strand. RNA primase attaches more RNA primer in the remaining gaps on the lagging strand and DNA polymerase extends the primed segments, yielding a series of Okazaki fragments. RNA primers are then removed by 5'-3' exonucleases (RNase), which DNA polymerase replaces with DNA nucleotides. A single gap in the sugar-phosphate backbone (nick) on the leading strand and several gaps in the sugar-phosphate backbone on the lagging strand can be found. DNA ligase fills these gaps in by forming a phosphodiester bond between a 5'-phosphate and a 3'-hydroxyl group, thus completing the newly replicated ‘daughter’ strands of DNA. Cell division can now occur.²³⁻²⁵

Overall, replication has a proofreading mechanism; DNA polymerases are in general extremely accurate, they can remove nucleotides from the end of a strand in order to correct mismatched bases. However, mismatches can persist; hence creating genetic mistakes (*see section 1.4.5 for further details*).

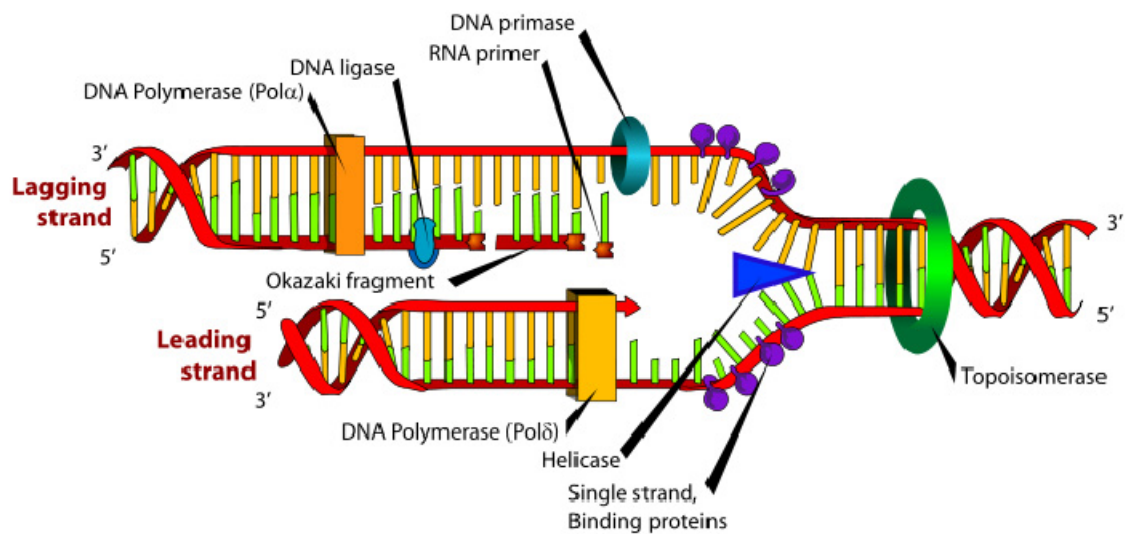


Fig. 1.14: DNA replication within the nucleus of a cell.

1.4.2 DNA Transcription²⁷

Transcription, or RNA synthesis, is the process of creating an equivalent RNA copy of a sequence of DNA, and is the first step leading to gene expression. The existence of base pairs provides a complementary language that can be converted back and forth between RNA and DNA in the presence of the correct enzymes.

The DNA helix unwinds and transcription starts at the promoter region. The promoter region is a particular location where the enzyme RNA polymerase initiates RNA synthesis. RNA polymerase reads the sequence of only the ‘template strand’ of the DNA duplex in a 3’ to 5’ manner, producing a complementary antiparallel RNA strand (where T is replaced by U). Transcription continues until one entire gene has been created where the RNA sequence (same sequence as the ‘coding strand’ of DNA) has been transcribed from the DNA sequence (transcription unit). The new RNA strand separates from the DNA. If the gene transcribed encodes for a protein, messenger RNA (mRNA) is created from pre-messenger RNA, which has been chopped by a process called RNA splicing to remove the introns (RNA that does not code for proteins). It is then used to create a protein *via* translation (*see below*). Alternatively, the transcribed gene may encode for ribosomal RNA (rRNA) or transfer RNA (tRNA), other components of the protein assembly process, or other ribozymes.

A DNA transcription unit encoding for a protein contains not only the sequence that will eventually be directly translated into the protein ('coding sequence'), but also 'regulatory sequences' that direct and regulate the synthesis of a particular protein. The 'regulatory sequences' upstream from the coding sequence is called the '5'-untranslated region' (5'-UTR), and the sequence downstream from the coding sequence is called the '3'-untranslated region' (3'-UTR). Transcription has some proof reading mechanisms, however, they are not as efficient as the controls for DNA replication.²⁷

1.4.3 RNA Translation²⁸

RNA translation is the process by which proteins are expressed. There exists 20 different amino acids, and these are the building blocks of proteins.

During translation, mRNA leaves the cell nucleus to enter the cytoplasm, and tRNA is replenished from the nucleus as it is depleted. Translation proceeds in four stages: (1) activation; (2) initiation; (3) elongation; and (4) termination. During activation, a condensation reaction occurs, and the correct amino acid is covalently bound at its carboxyl group to the 3'-hydroxyl group of the correct tRNA, forming an ester, with the elimination of water. Initiation involves the subunits of the ribosome binding to the 5'-end of mRNA with the help of initiation factors. The mRNA is decoded by the ribosome to produce a specific amino acid chain. The ribosome facilitates decoding by inducing the binding of tRNAs with complementary mRNAs. tRNAs are made up of anticodon sequences. Using Watson-Crick base pairing, the triplet of tRNA bases form hydrogen bonds to the complementary triplet of bases that make up the codon sequences in mRNA. The tRNAs carry specific amino acids that are chained together one at a time, into a polypeptide as the mRNA passes through and is 'read' by the ribosome. The ester at the 3'-end of the tRNA carrying the amino acid for extension of the protein chain, reacts with the α -amino group of the terminal amino acid of the growing protein chain to form a new peptide bond. The tRNAs one by one are then expelled from the ribosome. Termination of the polypeptide chain occurs when the A site (tRNA-binding site) of the ribosome faces a stop codon (UAA, UAG or UGA). No tRNA can recognise or bind to this codon; the stop codon has no anticodon. The stop codon induces the binding of a release factor protein that prompts the termination of protein synthesis, and disassembly of the entire ribosome/mRNA complex.²⁸

1.4.4 The Genetic Code and the Wobble Hypothesis

The genetic code is universal. There are 4 bases in RNA (A, G, C, U), Crick and Brenner *et al.*²⁹ proposed that there are 64 (4^3) possible triplet codes. Theoretically only 22 codes are required, one for each of the 20 naturally occurring amino acids, in addition to a start codon and a stop codon to indicate the beginning and end of a protein sequence. Many amino acids have several codes (degeneracy) hence all 64 possible triplet codes are used. No two amino acids have the same code, however, similar codon sequences are observed for amino acids whose side chains have similar physical or chemical properties.

Crick³⁰ proposed that the first two base pairs of the codon-anticodon pairing triplet, that codes for a specific amino acid, uses the standard base pairing (G·C, A·U and *vice versa*) rather strictly. However, there exists some 'wobble' (flexibility) in the pairing of the third base, in that there exists some variation (I·A, U·G, U·I *etc*; I represents inosine which is an analogue of guanosine, and lacks an amino group on the 2-position of the purine ring, see *fig. 1.15*). This existence of 'wobble' base pairs (*fig. 1.16*) means that a single base in the 5'-anticodon position of tRNA is able to recognise either of the pyrimidines or purines in mRNA as its 3'-codon base pair. This explains the degeneracy of the genetic code (e.g. the codons for the amino acid threonine are ACU, ACC, ACA and ACG). Not all triplet base combinations are possible. Wobble base pairs have different geometries to Watson-Crick pairs and are not as stable.

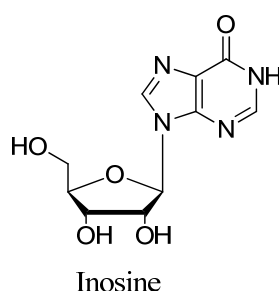


Fig. 1.15: The structure of Inosine.

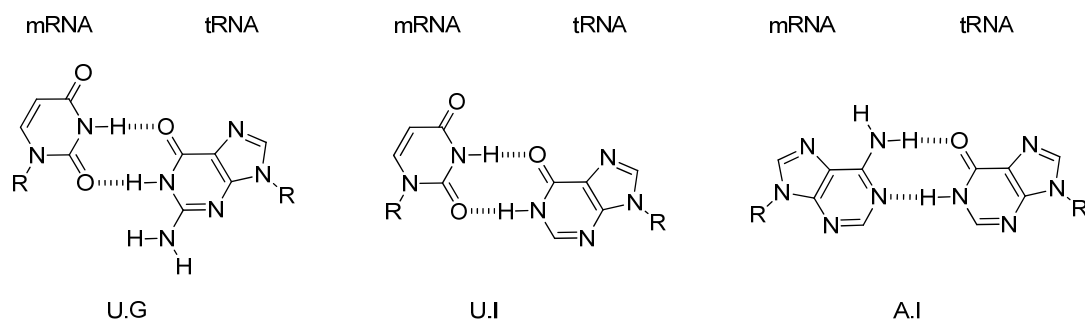


Fig. 1.16: The structures of a few ‘wobble’ base pairs; U·G, U·I and A·I respectively.

1.4.5 DNA Mutations and Repair³¹

It is estimated that without external influences, base pairing in DNA would allow more than 1 % incorporation of incorrect nucleotides. This is due to competition between the correct GC/AT base pairs and eight possible base mismatches; AA, GG, AG, CC, TT, CT, AC, GT. The first three of these are purine-purine mismatches, the second three are pyrimidine-pyrimidine mismatches and the final two are purine-pyrimidine mismatches. After a cycle of replication, purine-purine and pyrimidine-pyrimidine mismatches give rise to ‘transversion mutations’ and purine-pyrimidine mismatches produce ‘transition mutations’. DNA polymerase with their proofreading activity check for the incorporation of incorrect nucleotides, remove them and insert the correct nucleotide. However, the DNA polymerases cannot solely provide sufficient control and requires the aid of other proofreading and repair enzymes seeking out mismatches and effectively correcting them as the new DNA duplex is formed. These proof reading and repair enzymes that are used during DNA replication, work extensively with high efficiency, however, some mismatches do go uncorrected ($1 \text{ in } 10^9$) forming mutations. Mutations are changes in the DNA sequence of a cell’s genome. As well as occurring during DNA replication, mutations can arise from radiation, viruses, transposons (sequences of DNA that can move around to different positions within the genome of a single cell; process called transposition), and mutagenic chemicals. A single mismatch has the effect of providing an incorrect sequence for gene formation which in turn gives an incorrect RNA sequence by altering the code of the codon triplet. This then leads to the incorporation of one incorrect amino acid residue. Which in turn can lead to the expression of an incorrect protein, or a protein that possesses different functions to that of the protein that should have been expressed. In many cases when mismatches have

not been identified during replication, there are no disastrous consequences; some have no effect. However, if the DNA sequence is coding for a regulatory gene that controls the expression of an important protein, the presence of a mismatch can decrease the efficiency of regulation. This can lead to over production/under production of the specific protein, causing tissue to eventually degenerate, malfunction or develop cancers which can then lead to cell death. Occasionally however, the presence of a mismatch can be beneficial for example, by enhancing the region of the genome responsible for disease defense, hence making the host more resistant to attack from bacteria and viruses. Or, the mutation could produce a protein that works better than the original or performs a new function, hence the mutant organism will have a competitive advantage over its neighbours, and hence enabling it to spread its new gene through the population.³¹

The post-replicative DNA repair enzymes are able to distinguish the newly synthesised strand of the DNA duplex that requires repairing (the daughter strand) from the template strand (parental strand). Normally in the parental strands, the 5-position of some of the cytosine bases are methylated while the daughter strands are not. It is only after the proofreading and repair enzymes have checked that the daughter strand is a faithful reverse complementary copy of the parent, methylation of the specific cytosine bases in the daughter strand is carried out by an enzyme. The exact location within the DNA duplex that requires repairing is determined by considering the stability of the duplex from hydrogen bonding and base stacking interactions. It is these interactions that are important in the driving force for the formation of duplex DNA, which is an entropically unfavourable process. Prior to DNA duplex formation, the complex network of hydrogen bonds formed from the polar groups (amido, amidino, guanidino, carbonyl) of the heterocyclic bases in the single strands of DNA with the surrounding water molecules have to be broken. Overall, for duplex formation, the stability from the formation of the inter-base hydrogen bonds and base stacking interactions must be greater than the instability obtained from the loss of entropy and the energy required in breaking hydrogen bonds (endothermic process) from the DNA single strands to water. Mismatches (*see fig. 1.17, fig. 1.18 and fig. 1.19 for a few of the various kinds of mismatches*) can form wobble base pairs (these types of base pairs are formed in the ribosome and partly responsible for the degeneracy of the genetic code), or they can involve the base mispairs to be present in their tautomeric forms (some of these are

almost perfect mimics of Watson-Crick base pairs in overall shape), they can also involve ionized bases, or the pairing schemes can involve *anti-syn* isomerisation about the glycosidic bond. These various forms of mismatches generally destabilise the DNA duplex thermodynamically, due to the change in the shape and geometry of the mismatched pair, from that of Watson-Crick base pairs. The “mis-shapen” base pairs are unlikely to form stable stacking interactions within the DNA double helix, and in some cases hydrogen bonding of the heteroatoms to the surrounding water molecules is also inhibited, hence further destabilising the base pair. This destabilisation gives local melting and enhanced rates of the double helix opening. It is this phenomenon that is usually recognised by DNA repair enzymes. Some mismatches, however are harder to detect, and these can sometimes persist due to being unrecognised by the DNA repair enzymes.³¹ Aboul-Elä *et al.*³² observed a hierarchy of mismatch stabilities that can be expressed as $G \cdot T > G \cdot G > G \cdot A > C \cdot T > A \cdot A = T \cdot T > A \cdot C = C \cdot C$. It has been well documented^{33,34} however, that this hierarchy does not correlate with the efficiencies of correction of the different mismatches. $G \cdot T$ and $A \cdot C$ transition mismatches are usually well repaired due to these mismatches enhancing the rate of helix opening; while $G \cdot A$ and $C \cdot C$ transversion mismatches are generally poor substrates of mismatch correction.

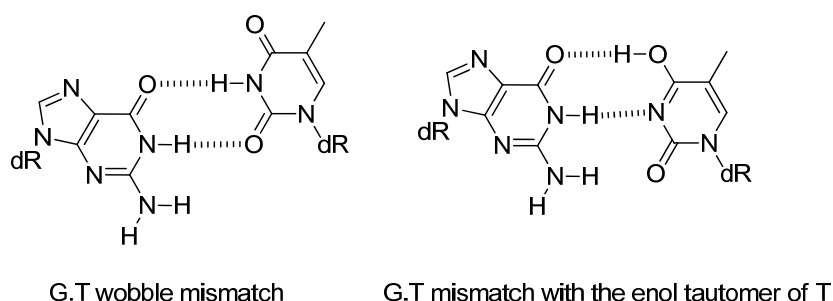


Fig. 1.17: The wobble $G \cdot T$ mismatch (left) and $G \cdot T$ mismatch with the enol tautomer of T (right).

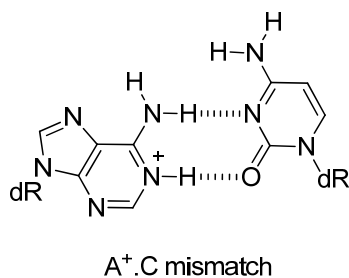


Fig. 1.18: The structure of the $A^+ \cdot C$ mismatch.

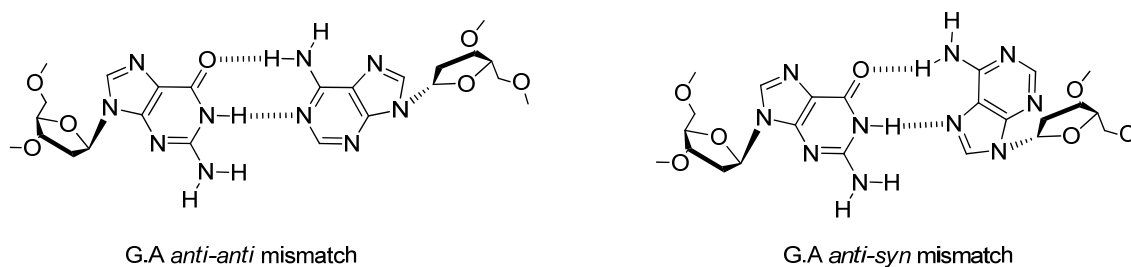


Fig. 1.20: The structures of the G·A *anti-anti* mismatch (left) and the G·A *anti-syn* mismatch (right).

1.4.6 Reverse Transcription

Reverse transcription was discovered independently by David Baltimore³⁵, Howard Temin³⁶ and Satoshi Mizutani³⁶. Reverse transcription occurs after an RNA retrovirus enters a host cell and with the aid of reverse transcriptase (an RNA-dependent DNA polymerase enzyme), single-stranded RNA of the retrovirus is transcribed into double stranded DNA. This is then integrated into the host chromosomal DNA and replicated along with it to produce the virus, reversing the direction of genetic transcription. Reverse transcription is extremely error-prone, due to reverse transcriptase enzyme having no proofreading ability, and it is during this step that mutations may occur; such mutations may cause drug resistance. HIV-1 is a well studied reverse transcriptase from human immunodeficiency virus (HIV) type 1.

Antiretroviral drugs^{37,38} such as zidovudine (AZT) (*fig. 1.21*) have been designed to block the HIV reverse transcriptase's enzymatic function and prevent completion of synthesis of the double-stranded viral DNA. Thereby preventing HIV from multiplying. Drugs like these are known as reverse transcriptase inhibitors (RTIs). RTIs are also used to treat tumours and cancer. RTIs can be subdivided into three groups: (1) Nucleoside analogue reverse transcriptase inhibitors (NRTIs); (2) Nucleotide analogue reverse transcriptase inhibitors (NtRTIs); and (3) Non-nucleoside reverse transcriptase inhibitors (NNRTIs). NRTIs and NtRTIs have essentially the same mode of action; they are analogues of the naturally occurring deoxynucleotides needed to synthesise the viral DNA, and they compete with the natural deoxynucleotides for incorporation into the growing viral DNA chain. However, unlike the natural deoxynucleotide substrates, NRTIs and NtRTIs lack a 3'-hydroxyl group on the deoxyribose moiety. As a result, following incorporation of an NRTI or an NtRTI, the next incoming

deoxyribonucleotide cannot form the next 5'-3' phosphodiester bond needed to extend the chain. Hence, the synthesis of the viral DNA is terminated. All NRTIs and NtRTIs are classified as competitive substrate inhibitors. In contrast, NNRTIs block reverse transcriptase by binding at a different site on the enzyme (a specific 'pocket' binding site) compared to NRTIs and NtRTIs. NNRTIs are not incorporated into the viral DNA but instead inhibit the movement of protein domains of reverse transcriptase that are needed to carry out the process of DNA synthesis. NNRTIs are therefore classified as non-competitive inhibitors of reverse transcriptase.

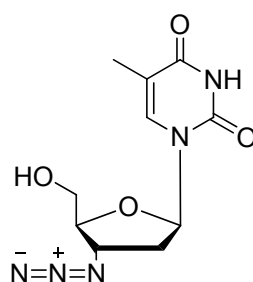


Fig. 1.21: The structure of Zidovudine/Azothymidine (AZT) a NRTI used in HIV treatment.

Although these reverse transcriptase inhibitors successfully prevent completion of synthesis of the viral double stranded DNA, it has proved difficult to develop compounds with the ability to entirely eliminate these life-threatening viruses from the host genome. Hence, it is for this reason that a number of oligonucleotide-based approaches are now being developed such as; antisense oligonucleotides (RNA interference, RNAi) and triplex forming oligonucleotides (TFOs), to target gene inhibition.

1.5 THE ANTISENSE APPROACH TO REVERSE TRANSCRIPTION AND GENE INHIBITION

The survival of viruses and bacteria within a host genome is dependent on the formation of their unique proteins. In the human genome, the mRNA sequences that encode these essential proteins are absent. Hence, methods that eliminate the foreign mRNA molecules without interfering with the mRNA of the host are potentially of great therapeutic value.

Antisense oligonucleotides are designed to hybridise to specific mRNAs and inhibit expression of the particular proteins that are for example, important for the replication of a virus, or for the uncontrolled growth of cancer cells. Elimination of the target mRNA is the eventual result of antisense. A hybrid duplex is initially formed between the mRNA and antisense oligonucleotide; the ribosome of the virus is therefore unable to assemble around the viral mRNA and hence cannot direct protein expression, thereby inhibiting the synthesis of viral proteins. Ribonuclease-H (RNase-H), which is present in normal cells, has endonuclease activity that enhances the antisense effect on mRNA molecules hybridised to antisense oligonucleotides, thereby destroying the mRNA strand. The antisense effect becomes catalytic. Hence, with the aid of RNase-H, a single antisense oligonucleotide can be involved in the destruction of many mRNA molecules.³⁹

The successful incorporation of antisense oligonucleotides into cells, without degradation by DNA enzymes, requires chemical modification. Modifications are normally carried out on the phosphodiester backbone of the antisense oligonucleotides. Modifying this moiety inhibits the action of 5'- and 3'-exonucleases, as well as endonuclease. DNA enzymes with 5'-exonuclease activity digest oligonucleotides from the 5'-end, 3'-exonucleases digest from the 3'-end, and endonucleases digest from within the DNA chain. Modifications can also be made to the sugar moiety to prevent oligonucleotide degradation by DNase enzymes. Such modifications can be made by adding functional groups to the 2'-position of the sugar; 2'-*O*-methyl; 2'-*O*-methoxyethyl; 2'-*O*-aminoethyl etc. It is possible to combine more than one modification in a single oligonucleotide.³⁹

The antisense approach to gene inhibition is very sophisticated however it is difficult to achieve complete inhibition of a specific mRNA when considering the vast numbers of mRNA molecules that are present within a cell. Importantly, feedback mechanisms exist that can lead to increased mRNA production in response to antisense inhibition of a particular message.

A more attractive approach to blocking the synthesis of specific proteins, rather than by antisense, is by the direct inhibition of DNA.^{40,41} In comparison to the thousands of copies of mRNA molecules present when considering the antisense approach, there are

only two copies of each DNA target site per diploid cell, hence, a better opportunity for the interacting agent to reach its specific target. The direct switching off of genes by an external agent in this way is an extremely attractive approach, and can be achieved in principle by sterically blocking the double helix so that proteins required for replication and transcription can no longer bind. The regions of DNA involved in the regulation of gene expression can be a target for inhibition; the formation of intermolecular DNA triple helices offers this possibility.

1.6 DNA TRIPLE HELICES AND THE ANTIGENE APPROACH TO GENE INHIBITION⁴²⁻⁴⁶

1.6.1 The Discovery of and an Introduction to DNA Triple Helices

Intermolecular DNA triple helices are formed when a third strand oligonucleotide (triplex-forming oligonucleotide, TFO) binds in the major groove of DNA, making specific contacts with substituents on the exposed faces of the base pairs (*fig. 1.22*).⁴³ The formation of DNA and RNA triple helices was first observed in 1957 by Felsenfeld *et al.*^{47,48} using x-ray fibre diffraction studies on mixing polyU and polyA in the ratio 2:1. Further studies showed that polyC and polyG can generate similar structures under conditions of low pH.⁴⁹ Triple helices are able to form as a result of the hydrogen bonding donors and acceptors located on the exposed faces of the purine bases of DNA within the major groove. The two additional hydrogen bonds are able to form between the duplex purine residues and the residues in the third strand; these bonds are Hoogsteen hydrogen bonds.⁴⁰ Duplexes capable of forming triple helices contain purine bases on one strand and pyrimidine bases in the complementary strand, hence a prerequisite for triplex formation is homopurine/homopyrimidine tracts within DNA.

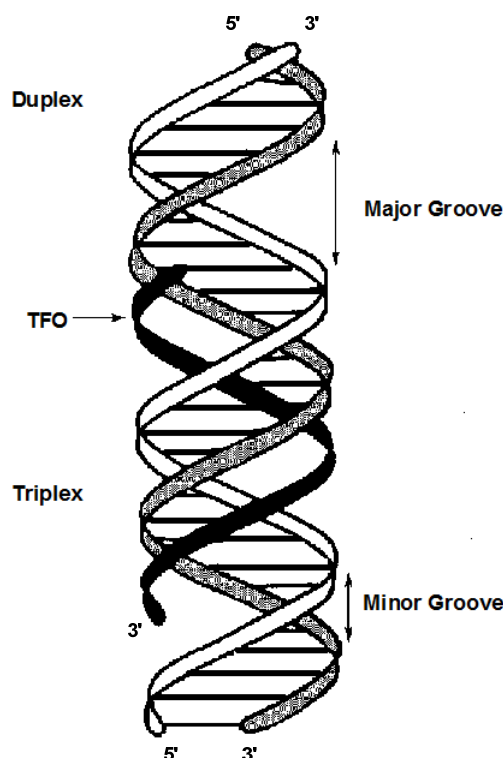


Fig. 1.22: Ribbon model demonstrating the relative position of a TFO (*black strand*) in the major groove of DNA. Watson-Crick base pairs are represented by the *solid black lines*. Hoogsteen hydrogen bonds are not shown.⁴³

The first biological role for triplexes was identified in 1968. Morgan and Wells⁵⁰ showed through an *in vitro* assay that transcription by an *E. coli* RNA polymerase could be inhibited by the presence of an RNA third strand. However, it was not until 1987 when it was discovered that the formation of DNA triple helices provided a means for designing DNA sequence specific agents.⁴⁰ Le Doan *et al.*⁴¹ demonstrated that a TFO containing all α -configured T bases with covalent attachment of an azidoproflavine derivative (intercalating mutagenic reagent; *fig. 1.23*), could form a triple helix with a stretch of A·T base pairs within the major groove of the DNA duplex. Hydrogen bonds were formed between the thymidine bases of the mutagen-activated TFO to the adenine bases within the major groove of the DNA duplex. The azidoproflavine derivate could then be excited by light to induce crosslinking reactions to either strand of the target duplex. These crosslinked species could then undergo cleavage reactions under alkaline conditions, thereby inducing irreversible damage within the target sequence, therefore, creating a potential method for controlling gene expression *in vivo*. Thus, TFOs not only have the potential to control biological processes such as transcription, DNA

replication, DNA repair and recombination, but to also act as a type of ‘molecular scissors’.

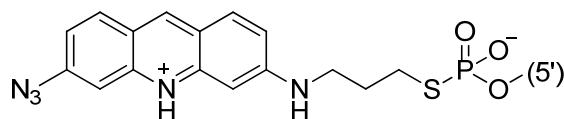


Fig. 1.23: The structure of the azidoproflavine derivative used in the DNA mutagenesis studies conducted by Le Doan *et al.*⁴¹

Naturally occurring intramolecular DNA triple helices exist. H-DNA^{46,51-56} forms when the third strand is provided by one of the strands of the same duplex DNA molecule at a homopurine-homopyrimidine mirror repeat sequence ((dT-dC)_n·(dA-dG)_n repeats) (*fig. 1.24*).⁵⁷ The H-DNA intramolecular triple helix is formed by the pyrimidine strand and half of the purine strand, leaving the other half of the purine strand single stranded. Transition to H-DNA occurs when the DNA is exposed to negative supercoiling (twisting of DNA in the opposite direction of the helix) and low pH (half of its C residues are protonated). The first polymorph of DNA of this kind was discovered in 1985 within a sequence (dA-dG)₁₆ in the polypurine strand of a recombinant plasmid pEJ4.⁵⁸ There exists many different isomeric forms of H-DNA whose structures differ depending on base sequence, sequence length, pH, ionic strength and topological stress. H-DNA motifs are abundant in the human genome, and it is thought that they have roles in DNA replication, transcription and recombination.



Fig. 1.24: The structure of the intramolecular triple helical DNA polymorph H-DNA.⁵⁷

TFOs can be used in genetic engineering. As a result of their ability to recognise sequences within double-helical DNAs with high specificity, TFOs constitute a valuable tool for site-directed modification of genomic DNA, and in the use for gene manipulation and gene function. Rather than designing molecules (agonists – a substance which fully activates the neuronal receptor that it attaches to; and antagonists – a substance that attaches to a receptor within a cell but does not activate it) to bind to specific receptor sites, TFOs offer an alternative pathway for drug design. Using molecules for drug design requires high specificity of the drug molecule with the target pharmacological receptor. This requires knowledge of the primary and secondary structures of the receptor site itself. The structure of DNA is known in precise molecular detail. Hence, the use of TFOs to target a specific region of DNA allows these sequence-specific agents to bind with high affinity and efficacy. TFOs are able to bind selectively to DNA, they are also able to adopt the same geometry as the helical duplex while binding to the major groove, turning through 360° every 10 base pairs. This is a result of being composed of the same repeat units as DNA.

1.6.2 Triplex motifs and their Structure

Depending on its base composition and binding orientation relative to its DNA target site, a TFO can be categorised into either the pyrimidine or purine motif. In the pyrimidine motif, the TFO is pyrimidine-rich (consisting mainly of C and T) and runs parallel to the duplex purine strand, the triplex is stabilised by the formation of Hoogsteen hydrogen base pairs⁵⁹, and characterised by T·AT and C⁺·GC triplets. These triplets are isomorphic (if the C-1' atoms of the Watson-Crick base pairs are superimposed, the positions of the C-1' atoms of the third strand are almost identical), this minimises backbone distortion of both the third strand and duplex between adjacent triplets. It is also possible to form a G·GC triplet within this motif, however it is not isomorphic with the T·AT and C⁺·GC triplets. Protonation at N³ of cytosine is essential for formation of Hoogsteen bonding with N⁷ of guanine which occurs only at low pH (< 6.0) (*fig. 1.25*).

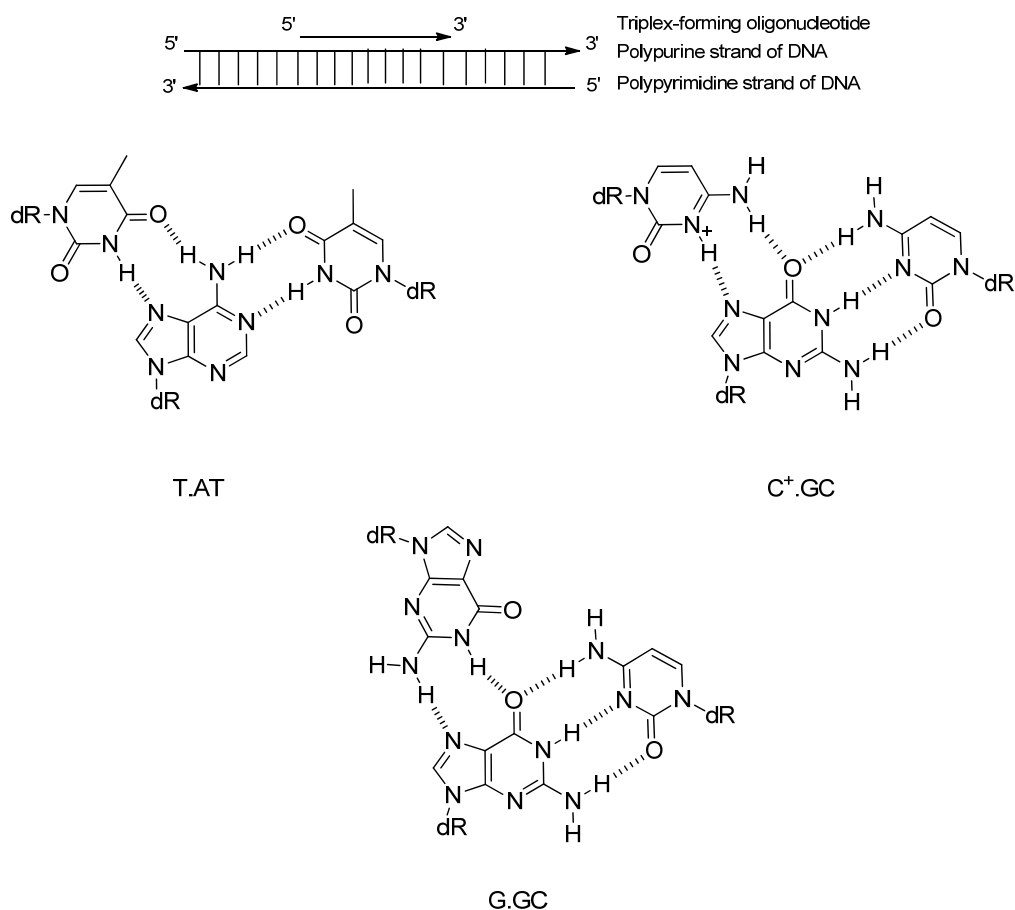


Fig. 1.25: The pyrimidine binding motif. *Above*, binding of a TFO in a parallel orientation to the polypurine strand of DNA. *Below*, structures of the parallel triplets T·AT, C⁺·GC and G·GC.

In the purine motif, the TFO is purine-rich (consisting mainly of A and G) and runs antiparallel to the duplex purine strand. Reverse-Hoogsteen base pairs enable reasonable stacking interactions for the triplex to be stabilised. The purine motif is characterised by A·AT, G·GC, and T·AT triplets (*fig. 1.26*).⁶⁰ These triplets are strictly not isomorphous as in the case for the pyrimidine motif.

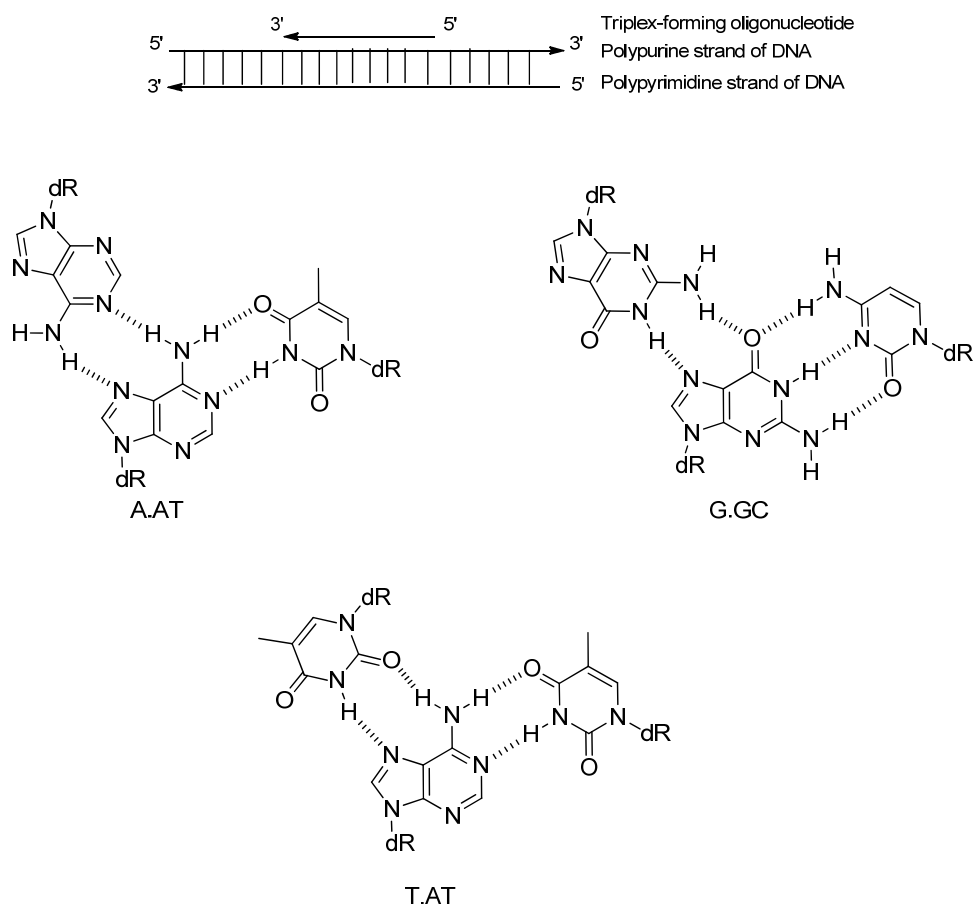


Fig. 1.26: The purine binding motif. *Above*, binding of a TFO to the polypurine strand of DNA in an antiparallel orientation. *Below*, structures of the antiparallel triplets A·AT, G·GC and T·AT.

G·GC and T·AT triplets can be formed in both the parallel and antiparallel motifs, hence, GT-containing oligonucleotides have great diversity, in that they can be designed to bind in either the parallel or antiparallel orientation. The number of GpT and TpG steps influences the stability of the triplex, and hence the binding orientation of the TFO. This is a result of backbone distortion imposed by the non-isomorphic nature of these triplets.⁶¹ In general, antiparallel triplexes are more versatile than parallel triplexes.⁴⁶

Each base pair (A·T, T·A, G·C and C·G) has a unique profile in the major groove. It is therefore possible to recognise specifically and selectively target any unique DNA sequence by probing the major groove through affinity of a base in a TFO, and its discrimination of a complementary base pair in the DNA duplex.

Intermolecular triple helix formation currently suffers from several limitations: (1) at all pH, triplexes in general form with lower stability than their duplex counterparts, this is partly due to the unfavourable bringing together of three polyanionic DNA strands in close proximity; (2) parallel triplexes are thermodynamically stable at room temperature below pH 6 in aqueous buffer, however less stable at physiological pH (pH 6.2 and above); (3) since there is no simple means of recognising pyrimidine bases, triplex formation is generally restricted to homopurine/homopyrimidine tracts.^{62,63}

1.6.3 Limitations to triple helix formation

1.6.3.1 Triplexes form with lower stability than their duplex counterparts

Triplex formation is governed by several factors including length, base composition, divalent cations, and temperature. It is kinetically very slow compared to the rate of duplex annealing.⁶⁴ In comparison to a duplex, there is a greater negative charge density within a triplex, hence greater repulsion between the three negatively charged phosphodiester backbones. In addition, the interactions generated by Hoogsteen hydrogen bonds are weaker than that generated by Watson-Crick hydrogen bonds. Also, triplex formation causes a conformational change in the underlying duplex, and this is also thought to contribute to slowing the process of triplex formation.⁶⁵

The polyamines; spermine, spermidine, and the diamine putrescine have been found to be present in mammalian cells in millimolar concentrations.⁶⁶ Their primary and secondary amino groups are all protonated at physiological pH and carry four, three and two positive charges, respectively. These polyamines are thought to play a key role in cell proliferation. They bind strongly to duplex DNA and stabilise its structure, neutralising the repulsive forces between the phosphate groups within DNA, and directly interacting with the DNA bases.⁶⁷⁻⁶⁹ At physiological concentration, the polyamines are able to interact with triplex DNA through electrostatic forces, and have shown to stabilise these structures too in the same way as for duplex DNA.⁷⁰ The stability arises through screening of the charge interactions between the three negatively charged phosphodiester backbones, allowing triplex DNA to form more readily.⁷⁰ The efficacy to increase triplex stability increases with their charge: spermine > spermidine > putrescine.⁷¹

The stability of triplexes in the purine motif is pH-independent, however depends on the presence of bivalent metal cations.⁵³ Although, only some bivalent cations are efficient in stabilising these triplexes, while others are not, and cation requirements are different for different sequences.⁷² In addition, physiological concentrations of monovalent cations, K^+ and Na^+ in particular, inhibit/destabilise these triplex motifs.^{73,74} Another requirement within this motif is that the purine TFOs must generally be guanine rich (> 65 %) to form stable triplexes.⁷⁵ However, for some G-rich TFOs, their binding is inhibited by the intracellular concentration of K^+ . They can be involved in auto-association processes via guanine-quartets that can compete with triplex formation. These G-quartets induce additional interactions with specific proteins, leading to side effects in cells.^{44,74,76}

The requirement for H^+ ions exists for the stability of triplexes in the pyrimidine motif, hence they are pH-dependent. These triplexes are also stabilised by the presence of divalent cations such as Mg^{2+} , Ca^{2+} and Zn^{2+} .⁴⁸

1.6.3.2 Parallel triplex formation suffers from the requirement for low pH

Parallel triplex formation generally requires conditions of low pH (< 6.0) necessary for protonation of N^3 of the third strand cytosine, to establish two hydrogen bonds with guanosine in the $C^+ \cdot GC$ triplet (*fig. 1.25*). However, for the use of TFOs in the pyrimidine motif for genetic manipulation, triplex formation is required to occur at physiological pH, at which the cytosine base will not be protonated.

The $C^+ \cdot GC$ triplet has been reported to provide an increased stability on a triplex relative to the $T \cdot AT$ triplet.⁷⁷⁻⁷⁹ This extra stability is not only attributed to the additional hydrogen bond, but also to the existence of stronger hydrogen bonds, favourable electrostatic interactions between the positive charge on a protonated cytosine and the negatively charged sugar-phosphate backbone, and favourable stacking interactions between the charged base and the π -stack.^{77,80} Without protonation, the $C \cdot GC$ triplet only contains one hydrogen bond between the exocyclic N^4 of cytosine and 6-keto group of guanine in GC; this triplet exerts a destabilising effect on all triplexes at neutral pH.⁸¹

The pK_a of deoxycytosine is ~ 4.3 , although this is increased on triplex formation due to the polyanionic environment, and is higher in the centre ($pK_a \sim 9.5$) than the termini ($pK_a \sim 6.2-7.2$) of a triplex,⁷⁷ optimum $C^+ \cdot GC$ triplet stability still requires acidic conditions. Terminal cytosine bases provide an overall less stabilising effect on the triplex, and this can be attributed to the poor stacking interactions that are encountered, with only one nearest neighbour at the 3'- or 5'-end.^{77,81} TFO sequences containing long contiguous runs of cytosine bases have shown to destabilise the triplex, this has been attributed to the electrostatic repulsion of adjacent protonated cytosines, and hence alternate CT base sequences are much preferred.^{75,78}

Several cytosine base analogues have been designed with the aim of forming Hoogsteen hydrogen bonds with guanine at physiological pH, these are discussed in section 1.6.4.2.

1.6.3.3 Triplex formation is restricted to homopurine/homopyrimidine tracts

Triplex formation requires the presence of oligopurine tracts on the target DNA. This limitation arises from the fact that within the major groove of duplex DNA, the purine bases have two vacant Hoogsteen hydrogen bonding sites (*fig. 1.27*); they can therefore form stable triplets. Pyrimidine bases having only one vacant site within the major groove of DNA, cannot form stable triplets, and would therefore not be efficiently bound by a TFO.⁶³

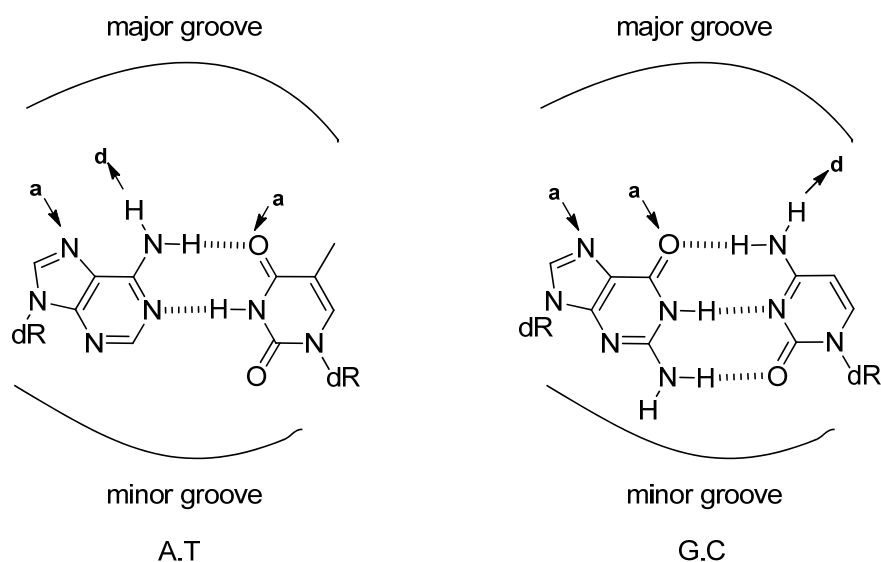


Fig. 1.27: The vacant Hoogsteen H-bonding sites (donor (**d**) and acceptor (**a**)) within the major groove of DNA for the Watson-Crick base pairs, A·T and G·C respectively. Purines have distinct binding profiles with bidentate vacant Hoogsteen sites whereas pyrimidines have only one.⁶³

However, within the genome it is often difficult to find sufficiently long uninterrupted homopurine sequences. Binding sites containing a single pyrimidine base can lead to potentially destabilising mismatches during triplex formation.⁷⁵

Along with the limitations discussed above, reagent delivery and cellular uptake are problems inherent to all oligonucleotide-based therapies.

1.6.4 Methods to Combat the Limitations to Triple Helix Formation^{44,61,63,75,80}

1.6.4.1 Increase Triplex Stability by Increasing the Strength of Triplex Binding

Several strategies to improve triple helix formation and increase its stability have been introduced and these are described below. They include: modifications to the sugar ring; the nucleobases, or; the phosphodiester backbone of naturally occurring nucleic acids.

1.6.4.1.1 Sugar Modifications⁸²⁻⁸⁴

On triplex formation, significant changes to the conformation of the purine strand of the underlying duplex occur which is unfavourable as it costs some of the intrinsic binding energy. For triplex formation, the optimum conformation of the sugars of the third strand is C^{3'}-*endo* (*N*) which imparts little distortion to the duplex purine strand, and hence preserves the binding energy.⁸⁴ RNA TFOs therefore, have a higher affinity for duplex DNA than DNA TFOs. Triplexes formed from DNA TFOs are the least stable type of triplex. The sugars within the TFO have a C^{2'}-*endo* (*S*) conformation, and conversion to *N*-type requires even more of the binding energy, which is unfavourable.

Oligonucleotide modifications in which the sugars are *N*-type are therefore able to produce more stable triplexes. The introduction of electronegative groups to the 2'-position of the sugar, as in RNA, strongly favours the *N*-type sugar pucker due to the gauche effect.^{85,86} The 2'-*O*-methyl modification has been shown to enhance triplex stability.^{65,87,88} The presence of this modification causes an increase in the hydrophobicity of the triplex, which effectively enhances the Hoogsteen bond stability. The modification occupies a narrow groove between the phosphate-sugar backbones of the parallel pyrimidine and purine strands that are in close proximity, thereby enhancing the rigidity of the triplex.⁶⁵ These results were also confirmed in NMR studies; the 2'-*O*-methyl group increases triplex stability by reducing the distortion of the duplex purine strand.⁸⁴ Furthermore, the addition of the methyl group to the 2'-hydroxyl significantly inhibits the action of nucleases in degrading the strand.⁸⁹

Locked nucleic acid (LNA) monomers (*fig. 1.28*), also called bridged nucleic acids (BNA) are conformationally restricted monomers, and adopt *N*-type sugar puckers. They were first developed by a number of groups between 1996⁹⁰ and 1997,^{91,92} and were called 'bicyclic nucleosides'. The most extensively reported of these consists of a 2'-*O*,4'-*C* methylene bridge which was developed independently by Singh *et al.*⁹³ in 1998 and Torigoe *et al.*⁹⁴ in 2001. These LNAs, when incorporated into TFOs displayed unprecedented hybridisation affinity towards complementary DNA and RNA. It was reported⁹⁴ that parallel triplex formation of LNA modified TFOs with duplex DNA produces a smaller negative (unfavourable) entropy change (ΔS) when compared to the entropy change associated with triplex formation of unmodified TFOs with duplex

DNA. This is due to a smaller entropic loss on triplex formation with the LNA modified TFOs at neutral pH, because these TFOs, in the free state are associated with an increased rigidity in comparison to the relatively ‘flexible’ unmodified TFOs in the free state. This therefore provides a favourable component to the ΔG (Gibb’s free energy change) and leads to an increase in the K_a (association constant). It is important to note here that fully modified LNA TFOs do not display binding towards dsDNA. It is thought that this is a result of their high rigidity, which is supported by structural studies by Gotfredsen *et al.*⁹⁵ There therefore exists a balance between the number of LNA residues in a TFO and the stability of the triplex formed.

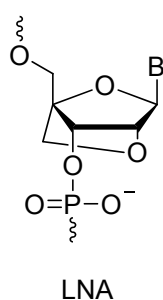


Fig. 1.28: The general structure of an LNA (locked nucleic acid) monomer; B represents any nucleobase.

1.6.4.1.2 Incorporation of Positive Charges

The stability of triplexes is limited by the charge repulsion that exists between the three negatively charged backbones that reside side by side. This is particularly so under physiological conditions of low ionic strength. Incorporating positively charged moieties into the TFO by either modifying the backbone, sugar or base, can alleviate this problem.

Positively charged backbone

The replacement of the phosphodiester linkages of the polyanion DNA with guanidine linkers to give the polycation deoxyribonucleic guanidine (DNG) (*fig. 1.29*) have been reported by Bruice^{96,97} and co-workers. Double and triple helices were formed between pentameric thymidyl DNG and negatively charged poly(dA), which displayed dramatic

stability according to their thermal denaturation temperatures.⁹⁶ The short DNG oligomer strand was reported to bind with unprecedented affinity and base-pair specificity.⁹⁶ In addition, at low ionic strength, these modified oligomers bound irreversibly to the target nucleic acid.⁹⁶ The synthesis of the ribose derivative RNG was also reported by Bruice⁹⁸ and co-workers. The guanidium linkages in RNG are neither susceptible to cellular nuclease, nor to chemical degradation under physiological conditions. Whereas, the labile nature of the phosphodiester backbone in RNA, causes it to be susceptible to nuclease hydrolysis, limiting its application. Hence, RNG is better suited as an antisense/antigene agent than RNA. RNG may also combine the favourable sugar conformation with the stabilising effect of the additional positive charge; it was reported that RNG-DNA forms very stable duplexes.⁹⁹ It was predicted that the overall structure of RNG is equilibrated as a B-DNA conformation, and the RNG strand adopts the general conformation of the DNA backbone. Hence RNG will bind specifically to B-DNA. Cationic dimethylaminopropyl phosphoramidate linkages (PNHDMAP), and *N,N*-diethyl-ethylenediamine linkages (DEED) (*fig. 1.29*) have also been reported¹⁰⁰⁻¹⁰² for the substitution of the anionic phosphate residues. A fully cationic PNHDMAP α -TFO was reported to form a highly stable parallel triplex that melts at a higher temperature than the duplex target. The recognition of the DNA duplex target with the cationic α -TFO was sequence-specific.¹⁰²

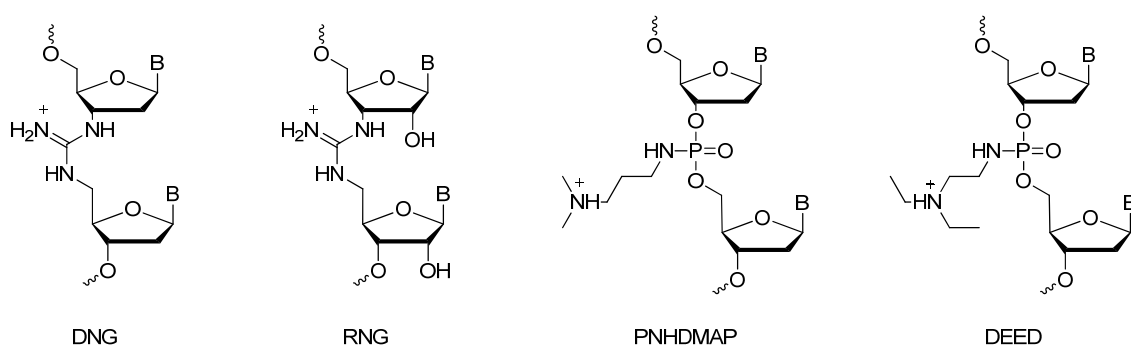


Fig. 1.29: The general structures of polynucleotides with cationic backbone modifications: (*from left to right*) deoxyribonucleic guanidine (DNG), ribonucleic guanidine (RNG), dimethylaminopropyl phosphoramidate (PNHDMAP), *N,N*-diethyl-ethylenediamine linkages (DEED).

Positively charged sugar

On triplex formation of single stranded RNA with duplex DNA, the 2'-hydroxyl groups of the RNA and the phosphate groups of the DNA second strand are in close proximity.¹⁰³ Cuenoud *et al.*¹⁰³ have reported the attachment of a short amino alkyl group (2'-*O*-aminoethyl) at the 2' position of thymidine units within a third strand (fig. 1.30); this allows the protonated amino group to form a specific intermolecular contact with a proximal phosphate group of the DNA duplex. Thermal denaturation studies showed that the stability of the triplex formed with this modified TFO was enhanced significantly in comparison to the triplex formed with the unmodified TFO which was displayed by the greater T_m value. This enhanced affinity of the TFO for the double stranded DNA could be down to the formation of specific contacts between the 2'-aminoethoxy side chains and the phosphate backbone of the DNA duplex, which is thought to contribute actively to the nucleation-zipping association process. Extension of the side chain by an additional methylene group, using a 2'-aminopropoxy group caused a significant decrease in triplex stability, suggesting that the 2'-aminoethoxy protonated group is ideally positioned to interact with a nearby phosphate group.

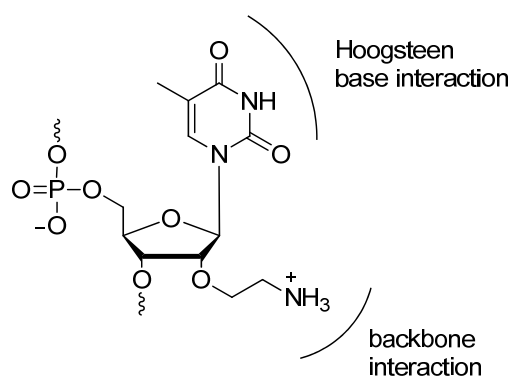


Fig. 1.30: The structure of the 2'-aminoethoxy-modified thymidine monomer.¹⁰³

Atsumi *et al.*¹⁰⁴ reported in 2002 the synthesis of 4'-(α)-*C*-aminoalkylpyrimidines and their incorporation into TFOs along with their triplex binding abilities. They found that TFOs containing 4'-(α)-*C*-(2-aminoethyl)-2'-deoxythymidine formed thermally stable triplexes with the complementary DNA target, in addition these TFOs were more resistant to nucleolytic digestion by DNase I endonuclease than unmodified TFOs. TFOs containing 4'-(α)-*C*-(2-aminoethyl)-2'-deoxy-5-methylcytosine were much more

resistant to 3'-exonuclease compared to the unmodified TFOs. Greater stabilisation was shown for the 2'-derivative and was attributed to the fact that the protonated 2'-aminoethoxy side chain is ideally positioned as a result of the gauche effect, to interact specifically with the nearby pro-R oxygen of the negatively charged phosphate group of the DNA second strand,¹⁰⁵ thereby forming a salt bridge.

Positively charged base

As a result of the positive charge in the C⁺·GC triplet, it is more stable than the T·AT triplet. Bijapur *et al.*¹⁰⁶ have replaced thymidines with 5-propargylamino-dU (U^P) (*fig. 1.31*) in the third strand TFO and studied its affinity for the duplex target site in comparison to the unmodified parent TFO, using UV melting and DNase I footprinting. The thermal denaturation studies for the triplex formed from the modified TFO showed a single UV-melting transition, which was > 20 °C higher than that of the parent triplex containing the unmodified TFO. A single melting temperature was observed at all pH, with the T_m decreasing with increasing pH; this is consistent with the requirement for protonation of the amino group. DNase I footprinting studies showed that triplexes formed with oligonucleotides in which the T was replaced with U^P displayed footprints at much lower concentration than their T-containing counterparts. In contrast to protonated C, adjacent U^P substitutions are not destabilising; the greater the number of U^P residues, the greater the stability of the TFO. The alkyl moiety of U^P also contributes to triplex stability by enhancing stacking interactions. Hence, Baijapu *et al.* demonstrated that the stability of DNA triplexes can be dramatically increased by using positively charged analogues of thymine, and that placing the charge in the major groove instead of the π -stacked bases, is a useful approach to stabilising triplexes.

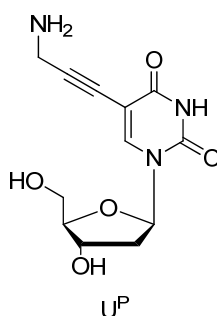


Fig. 1.31: The structure of 5-propargylamino-dU (U^P).¹⁰⁶

As discussed above triplexes are known to be stabilised by spermine.^{70,71} The covalent attachment of spermine at the C-5 position of uracil¹⁰⁷ and to the N4 position of 5-methylcytosine¹⁰⁸ (a naturally occurring cytosine derivative discussed in further detail in *section 1.6.4.2.1*) (*fig. 1.32*) both increased triplex stability at physiological pH. However, the complexes were reported to exhibit decreased selectivity.

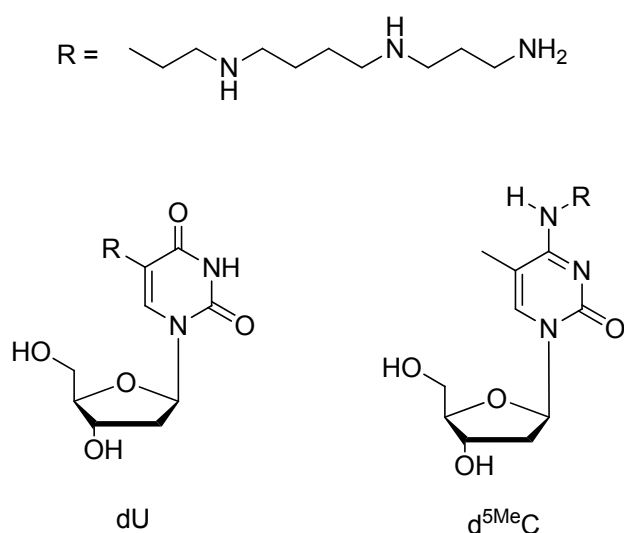


Fig. 1.32: The structures of the 5-spermine derivative of uracil (*left*) and the N4-spermine derivative of 5-methylcytosine (*right*). Where R = spermine (*top*).^{107,108}

It was surprising that the triplexes formed by the TFOs containing the methylcytosine derivative were stable at pH 7.3, even though N3 should be unprotonated at this pH. This suggests that the lack of the second hydrogen bond in the $\text{C}^+\cdot\text{GC}$ triplet can be compensated by favourable electrostatic interactions between the protonated spermine side chain and the phosphate backbone, and possibly hydrogen bonding interactions with the duplex bases. These triplexes were less stable at lower pHs. The substitution was more effective when placed at either end of the triplex than at the centre and caused a slight decrease in T_m with increasing the number of substitutions. These triplexes were less dependent on the presence of divalent metal ions such as Mg^{2+} .

Combining two modifications that individually are known to stabilise triplexes such as 5-aminopropargyl and 2'-O-aminoethyl should produce stable triplexes across a wide range of salt concentrations, particularly under physiological conditions.¹⁰⁹ Bis-amino dU (BAU)¹¹⁰⁻¹¹² (*fig. 1.33*) combines these two modifications and contains two positive

charges at physiological pH. It dramatically increases triplex stability when incorporated into TFOs as thymidine equivalents, with an increase in T_m of 7.7 °C per modification at pH 6.5,¹¹¹ and the stabilisation is much greater than when either modification is used alone. The two positive charges act in different ways to enhance triplex stability: the 2'-*O*-aminoethyl group interacts with a phosphate on the purine strand, while the 5-propargylamino group interacts with a third strand phosphate.¹¹⁰ This analogue is highly stabilising relative to thymidine and very selective for AT base pairs, with enhanced discrimination against pyrimidine inversions.¹¹¹ The high stability of the BAU·AT triplet permits triplex formation at physiological pH, even for sequences that contain several GC base pairs.

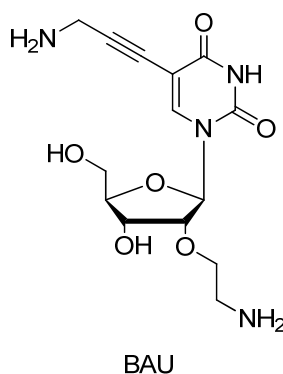


Fig. 1.33: The structure of bis-amino dU (BAU).^{110,111}

1.6.4.1.3 Modifications to the Backbone

Decreasing the charge repulsion between the three polyanionic DNA strands within a triplex is an attractive approach to increasing the affinity of a TFO to a duplex target. This therefore improves the stability of the triplex formed. TFOs that contain neutral backbones have been developed to do just this, and some examples are discussed below.

Triplex formation using short oligonucleotides containing alternating methylphosphonate and phosphodiester linkers (*fig. 1.34*) have shown to have important effects on triplex formation and to serve as substrates for nuclease enzymes.¹¹³ Both the phosphonate and phosphodiester linkages within the TFOs that were studied by Miller *et al.*¹¹³ were resistant to hydrolysis by spleen phosphodiesterase and by S_1 nuclease. The phosphodiester linkages of decamers were however cleaved slowly by snake venom

phosphodiesterase and by micrococcal nuclease. The rate of hydrolysis of the modified TFOs in comparison to the unmodified TFOs containing all phosphodiester linkages, were considerably slower. This difference in rate was thought to be due to the reduction of the negative charges which could reduce the affinity of the enzyme for the analogue, as well as non-ionic methylphosphonate groups being adjacent to the susceptible phosphodiester linkages. However, on using longer oligomers (19-mers) for triplex formation, it was reported¹¹⁴ that no triple helix was detectable by either T_m or gel analysis, when any of the chains A, T, or U was substituted by the corresponding methylphosphonate. The addition of a methylphosphonate oligomer strand to a normal DNA duplex also gave no detectable triplex. It was reported however, that the inversion of the anomeric configuration in the sugar moieties (from β to α) of methylphosphonate oligonucleotides can improve their poor binding to their nucleic acid targets.¹¹⁵ T_m values for hybrids formed with methylphosphonate α -analogues were obtained, and they were compared to those obtained for hybrids of the phosphodiester β - and α -oligomers. The non-ionic methylphosphonate α -oligonucleosides have a greater affinity for their complementary DNA and RNA strands than their homologues with natural β -anomeric configuration. With DNA as the target, the methylphosphonate α -oligomers formed duplexes more stable than the corresponding natural phosphodiester β -oligomers. Methylphosphonate α -dT₁₂ bound to its double stranded DNA target at low salt concentration of 0.1 M NaCl.¹¹⁵

Phosphoramidates (*fig. 1.34*) are 2'-deoxyribooligomers containing N3'→P5' internucleosidic linkages where O3' of the internucleoside phosphate is replaced by NH. They are similar to RNA in that they adopt an *N*-sugar puckering that favours Hoogsteen-type hydrogen bond formation with polypurine tracts in double stranded nucleic acids.¹¹⁶ Thermodynamic and kinetic analyses on triplexes in the pyrimidine motif have been carried out by Torigoe *et al.*¹¹⁷ Their results have indicated that at neutral pH, the phosphoramidate modification significantly increases the thermal stability of these triplex motifs (with double-stranded DNA as well as RNA·DNA heteroduplexes). In addition, at room temperature and neutral pH, increases in the binding constant by nearly two orders of magnitude have been observed for this triplex motif. Kinetic studies have demonstrated that the increase in binding constant (K_a) results mainly from a decreased dissociation rate constant (K_d).¹¹⁷ In further studies Torigoe *et al.*¹¹⁸ have combined this modification with a cationic copolymer. When

used on its own the cationic copolymer has been shown to increase the binding constant of triplex formation at neutral pH; this results from a considerable increase in the association rate constant rather than a decrease in the dissociation rate constant. Hence the combination of the two modifications has shown to cooperatively increase the binding constant for triplex formation at neutral pH, by four orders of magnitude. Kinetic complementarities between increased association rates by the copolymer and decreased dissociation rates by the phosphoramidate modification are the cause for the observed stabilisation. No negative interference was observed between these stabilising effects.¹¹⁸

Morpholino oligonucleotides are those in which the ribose sugar ring is replaced with a six membered morpholino ring and the phosphodiester linkage is replaced by a phosphorodiamidate (*fig. 1.34*). These oligonucleotides when compared to phosphodiester and phosphoramidate oligonucleotides form the most stable triplexes in the absence of Mg^{2+} , at a physiological monovalent cation concentration (low ionic strength). However, at neutral pH and in the presence of a high Mg^{2+} concentration, the phosphoramidate oligomers form the most stable triplexes.¹¹⁹

PNA (peptide nucleic acid) is an oligonucleotide mimic in which the entire deoxyribose phosphate backbone of DNA has been replaced by an uncharged backbone made up of 2-aminoethylglycine repeating units and the nucleobases are attached through a methylene carbonyl group (*fig. 1.34*).^{120,121} Mixed sequences of PNA containing all four natural nucleobases have been reported to be very potent DNA mimics. They form Watson-Crick base-paired duplexes with complementary DNA strands of high specificity and thermal stability.¹²² Homopyrimidine PNA oligomers form $PNA_2 \cdot DNA$ triplex structures with complementary homopurinic DNA sequences. These oligomers have been found to recognise double-stranded DNA targets by a mechanism that involves complete displacement of the pyrimidine DNA strand, leaving it looped out as a single strand.^{121,123} Thus indicating that the $PNA_2 \cdot DNA$ hybrid is more stable than the corresponding DNA triplex; this is a result of a much lower charge repulsion between the three strands. It has also been reported^{124,125} that a homopurine decamer made up of mixed A and G bases can bind to a single target in a double-stranded DNA by strand invasion and forming a Watson-Crick base-paired PNA-DNA duplex with the complementary DNA strand. The antiparallel duplex ($T_m = 69^\circ C$) is approximately

10 °C more stable than the parallel duplex ($T_m = 57$ °C), this result is comparable with thermal stability results obtained for PNA₂·DNA triplexes.¹²⁴ It was also reported that a mixed T/G PNA decamer was observed to bind to their Watson-Crick target sequences by invasion of the double-stranded DNA and displacing the identical DNA strand, forming new PNA-DNA complexes and P-loop formation, requiring two molecules of PNA, generating a 2:1 PNA:DNA triplex.¹²⁵ Cytosine-rich PNA oligomers form PNA-DNA₂ triplex structures; they add to their double-stranded targets as Hoogsteen strands.¹²⁵

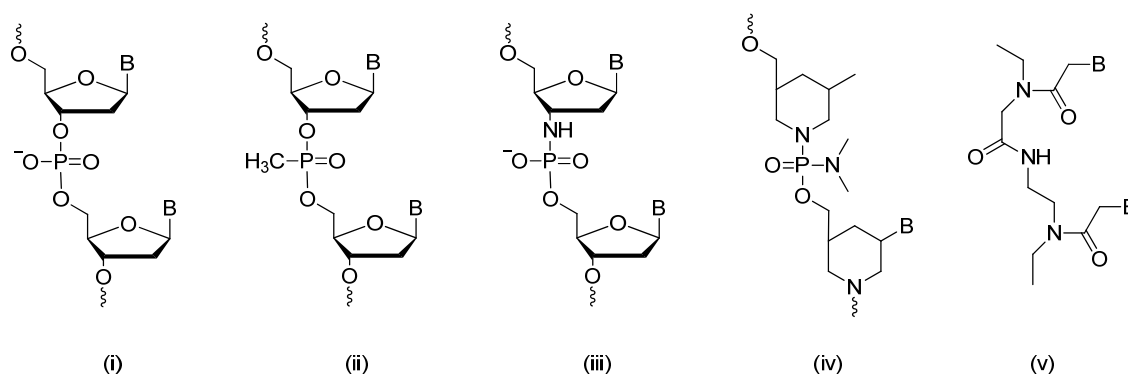


Fig. 1.34: Chemical structures of modified DNA backbones. *From left to right:* (i) phosphodiester, (ii) methylphosphonate, (iii) N3'-P5', (iv) morpholino and (v) PNA.^{113-119,121,125}

1.6.4.1.4 Improving Base Stacking Interactions

Base stacking (π - π interactions)^{22,126,127} involves the interaction between the aromatic moieties of the heterocyclic aromatic bases of DNA, and it affects the stability and structure of both duplex and triplex DNA. Extending the aromatic domain within a base can potentially increase third strand binding through stacking interactions. There have been many reported attempts that have extended the aromatic ring system of pyrimidine bases,¹²⁸⁻¹³⁰ by adding further aromatic rings across the 4-5^{128,129} or 5-6¹³⁰ positions, which should not affect the hydrogen bonding groups used in base pairing. However, triplexes containing these modifications did not show enhanced stability, hence the modification of pyrimidine nucleosides to form extended ring systems does not necessarily result in increased favourable base stacking interactions. The non-natural pyrido[2,3-*d*]pyrimidine F nucleoside can exist in two tautomeric forms F₁ and F₂ (*fig.*

1.35), and within triple helical complexes F mimics T and recognises AT base pairs (*fig. 1.36*) with a similar affinity to T, it exists in the F₂ tautomeric form.

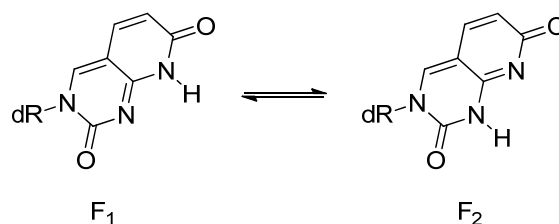


Fig. 1.35: The tautomeric forms of the F nucleoside.¹²⁸

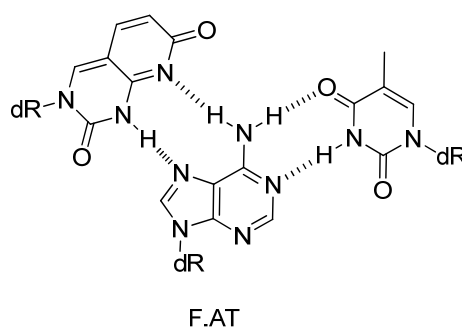


Fig. 1.36: The structure of the F·AT triplet.¹²⁸

Stacking energies can also be increased by the introduction of a hydrophobic substituent at the 5-position of a pyrimidine base, which concomitantly increases hydrophobic effects within the major groove. This is supported by the enhanced stability observed for the ^{Me}C⁺·GC triplet in comparison to the C⁺·GC triplet,^{81,131} and the stability of T·AT in comparison to U·AT. In the case of 5-methylcytosine, the enhanced triplex stability is partly due to an increase in its molecular polarizability with a concomitant increase of free energy for stacking, along with a source of positive entropy change due to the methyl groups filling a space in the major groove of the DNA duplex, causing a release of hydrating water molecules from the double helix to the bulk.¹³¹ It has been reported that the incorporation of the modified base 5-(1-propynyl)-2'-deoxyuridine (propynylU/pdU) (*fig. 1.37*) in the third strand of a triplex also leads to enhanced triplex stabilisation.^{132,133} The incorporation of pdU into TFOs increases the T_m for triplex melting by 2 °C per modification at 1 mM Mg²⁺ in comparison to the unmodified TFO. This indicates that pdU-modified TFOs are less dependent on Mg²⁺ concentration, than unmodified TFOs, and are capable of substantial binding even at physiologically low

Mg²⁺ concentrations that are close to intracellular concentrations.¹³⁴ The increased stabilisation of this modification is partly down to the extended aromatic cloud of the pdU nucleotide stacking well over the 5'-neighbouring nucleotides. Overall, the increases in base-stacking interactions and local hydrophobicity caused by the propynyl substituent, creates stable triplexes.¹³³ The incorporation of 5-(1-propynyl)-2'-deoxycytidine, pdC (*fig. 1.37*) into TFOs for triple helix formation creates a destabilising effect on the triple-helical complex due to a decrease in the pK_a of N3 of the heterocycle.¹³⁵

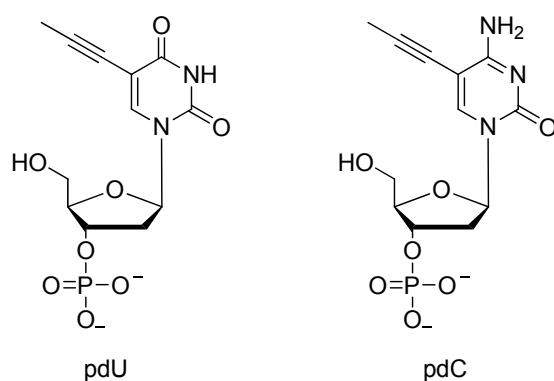


Fig. 1.37: The structures of pdU (*left*) and pdC (*right*).¹³⁵

1.6.4.1.4 The Incorporation of Triplex Binding Ligands

The stability of DNA triplexes can be enhanced by the incorporation of small molecules that are intercalators which can preferentially bind to triplex DNA over duplex DNA. A benzo[*e*]pyridoindole derivative (BePI) (*fig. 1.38*), which is composed of aromatic rings containing heteroatoms for stacking (intercalation) between the base triplets, and incorporates a positive charge to partially alleviate the triplex charge repulsion, has been reported to display the preferential triplex binding.¹³⁶ Upon irradiation with near-ultraviolet light, BePI induces covalent modifications of both Watson-Crick strands of the triplex, and it increases the thermal stability by more than 20 °C, hence strongly stabilising the triplex structure. In the absence of the third oligonucleotide, the strong and specific binding sites of the intercalator disappear, and this demonstrates a synergistic action of the intercalator and the TFO. The stabilising effect of BePI was shown to be much larger than that of the polycation spermine. A requirement for triple helix stabilisation by BePI is a long stretch of contiguous T·AT triplets due to the high

density of negative charges present at these stretches, whereas the presence of a positive charge in a C⁺·GC triplet hampers the binding of the cationic BePI.¹³⁶

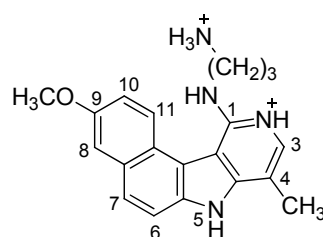


Fig. 1.38: The structure of the benzo[*e*]pyridoindole derivative with OCH₃ at the 9-position (BePI).¹³⁶

Other DNA binding ligands (acridine¹³⁷, triplex binding ligands¹³⁸, minor groove binding ligands¹³⁹ and daunomycin¹⁴⁰) can be covalently attached to a TFO to enhance triplex stability. The TFO enables sequence selectivity, while the non-specific DNA binding ligand can be used to enhance the affinity of the corresponding TFO. Triplexes that are formed at sites containing pyrimidine interruptions can also use this methodology to enhance stability.¹⁴¹ Psoralen (*fig. 1.39; see section 1.8 for further details on psoralen*), a photoreagent can also be added to a TFO to enhance triplex stability, it generates cross-links upon UV irradiation.¹⁴²

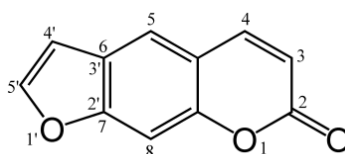


Fig. 1.39: The structure and numbering system used for psoralen.

1.6.4.2 Strategies to Combat the Dependency of Parallel Triplex Formation for Low pH.

It can be clearly seen that the recognition of guanine requires a structure presenting two hydrogen bond donors to interact with its O6 and N7 (see *fig. 1.1*). It has been previously discussed that several cytosine base analogues have been designed with the aim of forming Hoogsteen hydrogen bonds with guanine at physiological pHs and these analogues will be discussed here. These analogues can be divided into two classes

which are based on either a pyrimidine or a purine ring. These modifications can also be used in combination with sugar or backbone modifications.

1.6.4.2.1 Pyrimidine Base Analogues

5-methylcytosine ^{5Me}C^{80,131,143-146}

5-methylcytosine (^{5Me}C/^{Me}C) (*fig. 1.40*) is a naturally occurring base and its effect on triple-helix formation was first observed in polynucleotide structures.¹⁴³ It was found that triplex formation was possible at near pH 8 for the triplexes (poly(dTd^{Me}C)·poly(dGdA)·poly(dTd^{Me}C)) in which all cytosines were substituted with ^{Me}C. In comparison, the unmodified triplexes containing the unmethylated analogue required lower pH for triplex formation.¹⁴³ Thermal denaturation studies displayed higher thermal midpoint temperatures (*T_m*) for triplexes containing 5-methyl-2'-deoxycytosine in comparison to those containing 2'-deoxycytosine at pH 8.0. Hence, cytosine methylation strongly enhances triplex stability. These results and those obtained in studies by others^{131,144} demonstrated that the substitution of 2'-deoxycytosine with ^{Me}C extends the pH range at which triple helix formation can occur, up to near neutrality. However, not all triplexes containing this base are formed under physiological conditions.¹⁴⁵ The differences in the *pK_a* of 2'-deoxycytosine (*pK_a* ~ 4.3) and 5-methyl-2'-deoxycytosine (*pK_a* ~ 4.5) is measured to be ≤ 0.2 unit,¹⁴⁷ this slight increase in *pK_a* can be used to partly explain the extended pH range for triplex formation, however, it cannot affect the protonation at N-3 significantly. Methylation of C5 in ^{Me}C is thought to impart a positive entropic factor that enhances the stability of a triplex. The methyl group fills a space in the major groove of a triplex, causing a release of hydrating water molecules from the double helix to the bulk, thereby causing a positive entropy change.¹³¹ In addition to these effects, another likely source of the stabilising influence of the C5-methyl group is from increased base stacking interactions; methylation of the stacked bases increases their interaction energies.¹⁴⁶ Hence, the replacement of cytosine by ^{Me}C increases the stability of triplexes but does not alleviate the pH dependence.¹⁴⁷

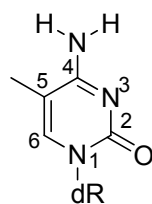


Fig. 1.40: The structure and numbering system of 5-methylcytosine (^{Me}C) deoxyribonucleoside.

*Pseudoisocytosine (isoC/^ΨC)*¹⁴⁸⁻¹⁵⁰

Pseudoisocytosine (^ΨC) (*fig. 1.41*) is a C-nucleoside (*discussed in further depth in chapter 2*). It is a non-charged cytosine analogue that is able to form Hoogsteen hydrogen bondings with guanine at neutral and basic conditions. This is because the N-3 position of ^ΨC is already bonded to hydrogen, hence being able to form the ^ΨC·GC triad, without any pH dependence. It has been reported that a 2'-methoxy derivative of ^ΨC which was incorporated into a third strand TFO stabilised triplex formation without pH-dependence.¹⁴⁸ ^ΨC overcomes the limitation presented with duplexes containing clusters of guanine residues, due to there being no existence of positive charge-positive charge repulsion along the third strand which is experienced when protonated ^{Me}C or dC are present in the third strand.¹⁴⁹ Triplexes containing the 2'-O-methyl derivative are more stable than those formed by the corresponding deoxyribooligonucleotides.¹⁴⁹ The incorporation of ^ΨC into bis-PNAs (two PNA segments connected in a continuous synthesis *via* a flexible linker) has been reported, they form triple stranded complexes with DNA ((PNA)₂/DNA), and display high thermal stability.¹⁵¹

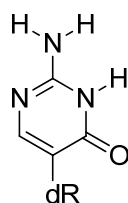


Fig. 1.41: The structure of pseudoisocytosine (^ΨC) deoxyribonucleoside.

6-Oxocytosine (^{Oxo}C)

6-oxocytidine (*fig. 1.42*), its 5-methyl derivative and its 2'-O-methyl derivative have been incorporated into TFOs and reported to form triplexes, recognising GC base pairs,

in a nearly pH independent (6.0-8.0) fashion.^{152,153} The methyl derivative encompasses slight destabilisation. Both Berresem *et al.*¹⁵² and Xiang *et al.*¹⁵³ have reported however, that at lower pH (< 6.5) cytosine/methylcytosine containing TFOs form more stable triplexes than those containing 6-oxocytosine/5-methyl-6-oxocytosine. It is thought that more favourable base-stacking interactions occur between the electron-rich cytosine/methylcytosine base and the adjacent bases, while the electron-poor heterocycles of the 6-oxocytosine and its methyl derivative causes weaker base stacking interactions with adjacent bases, and thus, decreases the stability of the triple helix. At low pH this effect is not compensated by more stable hydrogen bonds, which are present in 6-oxocytidine and its methyl derivative. The acidity at the 3-NH and 4-NH₂ groups is increased due to the presence of the electron-withdrawing oxo group, hence stabilising the hydrogen bonds to guanine.¹⁵² It was also thought that as cytosine and methylcytosine form one hydrogen bond in the triplex that contains a charged hydrogen-bonding partner; this enhances triplex stability, whereas 6-oxocytosine residues form only hydrogen bonds involving uncharged partners. At higher pH (physiological pH) however, the extent of protonation of cytidine/methylcytidine decreases, and triplexes containing 6-oxocytidine/5-methyl-6-oxocytidine, with neutral hydrogen bonds dominates in stability; 6-oxocytidine/5-methyl-6-oxocytidine becomes superior to cytidine/methylcytidine. Duplexes containing series of contiguous G·C base pairs cannot be effectively targeted with TFOs containing 6-oxocytidine/5-methyl-6-oxocytidine, as a result of ineffective base stacking between the adjacent residues and undesirable steric effects, due to the additional carbonyl at the 6 position.¹⁵⁴ Xiang *et al.*¹⁵⁵ reported the use of an acyclic 6-oxocytosine nucleoside derivative in which contains an acyclic or ring-opened carbohydrate residue, in TFOs, providing greater flexibility to overcome the undesirable steric effects observed when targeting contiguous G·C base pairs. However, as the number of acyclic linkers in the TFO was increased from two to six, triplex stability was observed to decrease slightly. This observation is possibly due to the flexible linker having to adopt the conformation most appropriate for triplex formation causing an unfavourable loss in entropy. This cancels any stabilising contribution from the enhanced base stacking.

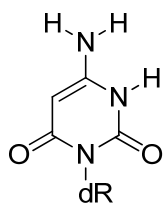


Fig. 1.42: The structure of 6-oxocytosine (^{oxo}C) deoxyribonucleoside.

2-Aminopyridine (P)

2-amino-5-(2'-deoxy-β-D-ribofuranosyl)pyridine (**P**)¹⁵⁶⁻¹⁵⁸ (*fig. 1.43*) is a C-nucleoside whose structure is similar to 2'-deoxycytosine, it is however considerably more basic, with a pK_a of 5.93 (compared with 4.3 for dC), which is closer to physiological pH than cytosine or 5-methylcytosine. This analogue was incorporated into 5'-psoralen linked oligonucleotides to form triplexes with a target duplex, the oligonucleotides were found to have considerably higher binding affinities for the target than the corresponding oligonucleotides containing either cytosine or 5-methylcytosine.¹⁵⁶ It has also been reported that oligonucleotides containing **P** bind to their duplex target with significantly greater affinity at higher pH than those containing either cytosine or ^{Me}C, hence relieving some of the pH-dependency. This therefore demonstrates that by enhancing the degree of N-3 protonation which is obtained through increasing the pK_a , the binding of cytosine analogues to G·C base pairs is best improved.¹⁵⁷ This increased stability is also evident at low pHs. TFOs containing **P** also form stable complexes with targets consisting of contiguous G·C base pairs.¹⁵⁷ The methyl derivative of **P**; 3-methyl-2-amino-5-(2'-deoxy-β-D-ribofuranosyl)pyridine (^{Me}**P**) (*fig. 1.43*) and the 2'-O-methyl derivative (**P**_{OMe}) (*fig. 1.43*) have also been reported to be incorporated into TFOs to form triplexes with target duplexes.¹⁵⁸ The replacement of cytosine or ^{Me}C residues with ^{Me}**P**, also results in an enhanced affinity of the TFOs to the duplex targets over a pH range 6.0-8.0, hence causing an increase in triplex stability and much less expressed pH dependence. The replacement of **P** by ^{Me}**P** (which is 0.4 pK_a units more basic than **P**) results in no further enhancement of triple helix stability at neutral pH. However, at pH values > 8.0, the effect of the pK_a difference between ^{Me}**P** and **P** becomes operative. A distinct loss of triplex stability has been reported when replacing **P** by **P**_{OMe} in the third strand, although there seems to be no relevant difference in the pK_a of these two

nucleosides.¹⁵⁸ This destabilisation may be caused by conformational differences in the sugar ring.

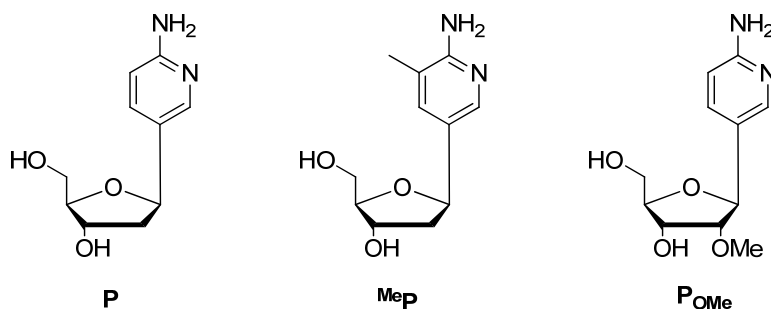


Fig. 1.43: The structures and numbering system of 2-aminopyridine (**P**), 3-methyl-2-aminopyridine (**^{Me}P**) and 2'-O-methyl-2-aminopyridine (**P_{OMe}**) deoxyribonucleosides.

1.6.4.2.2 Purine Base Analogues

Purine analogues can also be used for GC recognition, these are discussed below. The triplets that are formed from these analogues however, are not isostructural with T·AT, this is because the backbone is in a different position. Hence, equivalent analogues of T for recognition of A·T pairs may need to be used in combination.⁸⁰

8-oxo-adenine

Miller *et al.*¹⁵⁹ have reported that homopyrimidine TFOs containing the modified purine 8-oxoadenine are capable of forming stable triplexes with G·C base pairs (*fig. 1.44*) in target duplexes at neutral and basic pH. The keto form of the analogue is required for it to exist in the *syn* conformation. This conformation allows for the formation of two hydrogen bonds to G in the target duplex. One hydrogen bond forms between the N-7 proton of 8-oxoadenine with N-7 of guanine, and the other forms between the 6-amino protons of 8-oxoadenine with O-6 of guanine in the G·C base pair. In contrast to cytosine, protonation of the base is not required.

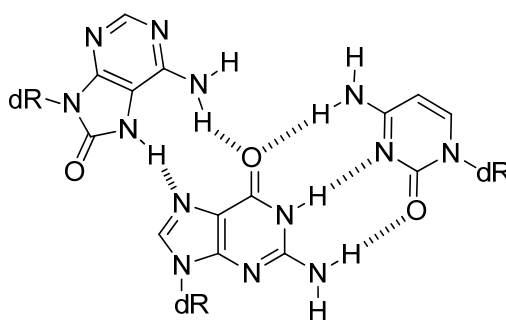


Fig. 1.44: The structure of the triplet formed between 8-oxo-2'-deoxyadenine and the G·C base pair.

The N⁶-methyl derivative of the 8-oxoadenine analogue (*fig. 1.45*), when incorporated into TFOs, has also displayed triplex forming ability to G·C base pairs in target duplexes. The triplexes form in a pH-independent fashion within the physiological range, without compromising specificity.¹⁶⁰ The increased affinity of this purine analogue for duplex DNA under physiological salt conditions has been found to be superior to ^{Me}C.

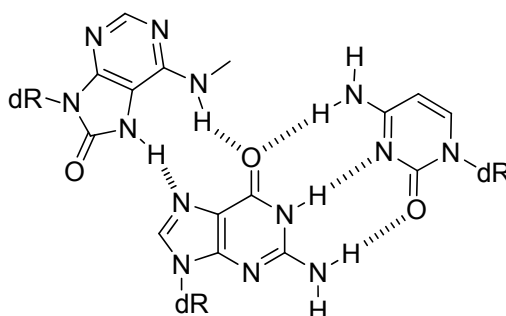


Fig. 1.45: The structure of the triplet formed between N⁶-methyl-8-oxo-2'-deoxyadenine and the G·C base pair.

These analogues recognise G·C and generate triplexes that have the same stability as those containing ^{Me}C at low pH. They can also be used for targeting contiguous guanines, which cannot be targeted with cytosine-containing oligonucleotides. However, they encounter reduced affinity at isolated Gs, as the triplets are not isomorphic with T·AT.⁶¹

It has been reported that an N⁷-glycosylated guanine, when incorporated into a pyrimidine oligonucleotide, binds with remarkable specificity to the G·C base pairs within a duplex forming the N⁷-G·GC triplet (*fig. 1.46*).¹⁶¹ The third strand orientation in the N⁷-G·GC triplet is parallel to the purine Watson-Crick strand; this is achieved by attaching the deoxyribose moiety to the N⁷-position of a guanine base. Hence, the Hoogsteen face of the base mimics the arrangement of hydrogen bond donors seen in protonated cytosine. At pH 7, N⁷-G binds to its target Watson-Crick G·C base pair with an affinity equal to that shown by MeC. However, at lower pH (5.2) this modified nucleoside analogue is not as stabilising as C⁺.¹⁶² This comparatively destabilising effect of the N⁷-G·GC triplet is thought to be due to the localised effect on base stacking interactions and the phosphodiester backbone; there exists slight displacements of all residues in the Watson-Crick strands. At pH 7, destabilisation as a result of the difficulty of protonating the C⁺·GC triplet outweighs any destabilisation of the triplex caused by the inclusion of N⁷-G.

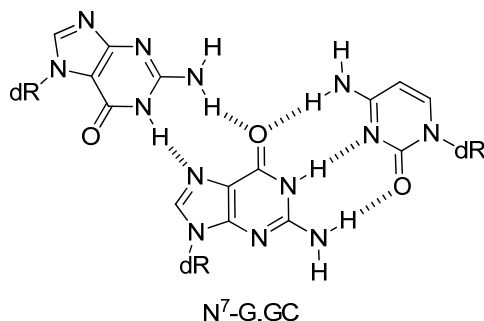


Fig. 1.46: The structure of the N⁷-G·GC triplet.

Other N⁷-glycosylated purine analogues include P1¹⁴⁵ (*fig. 1.47*) and N⁷-inosine^{163,164} (*fig. 1.48*). P1 binds G·C base pairs within a pyrimidine motif triplex similar to cytosine in terms of selectivity and strength, however, over an extended pH range. In addition, oligonucleotides containing P1 bind G-rich purine tracts in double helical DNA at neutral pH. The N⁸-glycosylated derivative of P1 generates P2 (*fig. 1.47*), however no triplex formation was observed using this analogue. Hence, the marked difference between the isomeric nucleosides P1 and P2 demonstrates the importance of the

positioning of the sugar-phosphate backbone along with the glycosidic conformation, for recognition of G·C base pairs.¹⁴⁵

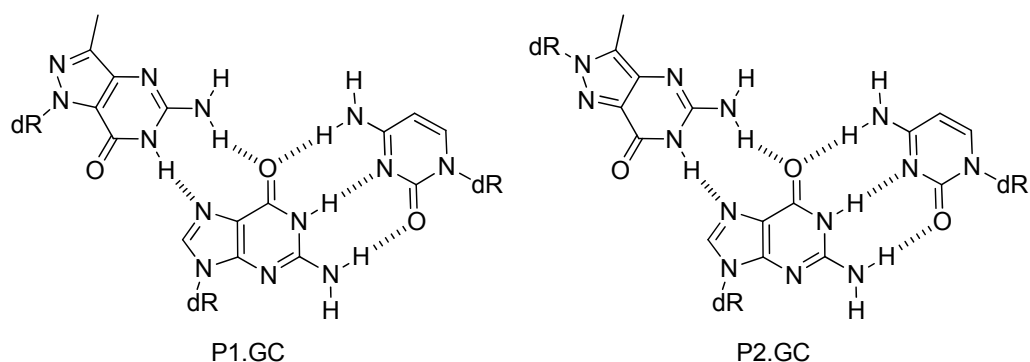


Fig. 1.47: The structures of the triplets P1·GC and P2·GC.

Recognition of the G·C base pair has also been demonstrated by N⁷-inosine (⁷H) and its derivatives *via* only one hydrogen bond between the N-H of inosine and N⁷ of guanine (*fig. 1.48*).¹⁶⁴ These analogues were incorporated into TFOs containing ^{Me}C, and displayed binding to a DNA target with equal to slightly increased stability compared to a control TFO in which the ⁷H residues were replaced by ^{Me}C. Also, TFO sequences containing multiply substituted ⁷H residues either in an isolated or in a contiguous manner still formed triplexes with their targets of comparable stability as the control sequences containing ^{Me}C, at pH 7.0. Hence, these results show that monodentate recognition in the ⁷H case to G·C base pairs, is energetically comparable to bidentate recognition in the ^{Me}C or N⁷-G cases. Although forming only one hydrogen bond to G·C base pairs, the stability of ⁷H can be attributed to the additional energetic contribution that may arise from enhanced intrastrand stacking around the modified bases alongside favourable electrostatic interactions between CH-2 of ⁷H and O-6 of guanine.¹⁶⁴ This interaction provides a small, positive, direct electrostatic contribution, which is energetically beneficial. In addition, this interaction alleviates any destabilising interaction that arises if a hydrogen bond acceptor is inactive or involved in secondary hydrogen bonding to solvent for example.¹⁶³

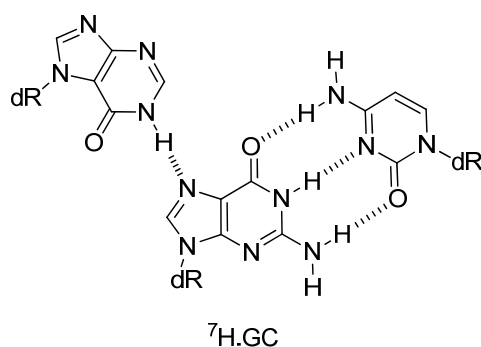


Fig. 1.48: The structure of the ${}^7\text{H}\cdot\text{GC}$ triplet.

1.6.4.3 Strategies to Combat the Requirement for Homopurine/Homopyrimidine Tracts^{112,165}

In the major groove of DNA, pyrimidine bases are presented with only one hydrogen bonding site on their Hoogsteen face. Hence triplets formed with the pyrimidine bases of a target duplex, would form with low affinity and reduced selectivity.^{112,165} Third strand recognition of cytosine and thymidine bases is therefore hard to achieve, restricting triplex formation to homopurine·homopyrimidine tracts. Described below are some of the strategies which have been employed for recognising pyrimidine interruptions using either natural bases or synthetic base analogues.

Natural Bases

The best combinations for recognising T·A and C·G base pairs using natural bases are G·TA (*fig. 1.49*) and T·CG (*fig. 1.50*).^{166,167} However, these triplets are much less stable than T·AT and C⁺·GC triplets as a result of local changes in the structure of the third strand, and multiple inversions are strongly destabilising; a 30-fold decrease in third strand affinity is associated with each addition of a pyrimidine residue.¹⁶⁸ The T·CG triplet can be accommodated within both parallel and antiparallel triplexes, while G·TA is limited to parallel structures.

The hydrogen bond in the G·TA triplet is able to form between the 2-amino proton of the third strand guanine and the O-2 group of thymidine within the duplex. For this hydrogen bond to exist the guanine base must adopt an anti conformation. The 5-methyl

group of thymine in the Watson-Crick base pair is thought to be involved in favourable stacking interactions with the third strand guanine, which increases the overall stability of the G·TA triplet. If a G·TA triplet has a 3'-adjacent T·AT triplet, an additional hydrogen bond can be formed between the unused 2-amino proton of guanine and the O-4 of the duplex thymine.^{165,169} Hence, G·TA is more stable when flanked by T·AT rather than C⁺·GC. In order to facilitate the structural changes associated when G is incorporated into a pyrimidine-rich TFO, a C^{2'}-*endo* to C^{3'}-*endo* sugar pucker conversion occurs for the guanine nucleoside, this positions the sugar in a more favourable position within the backbone.¹⁶⁵

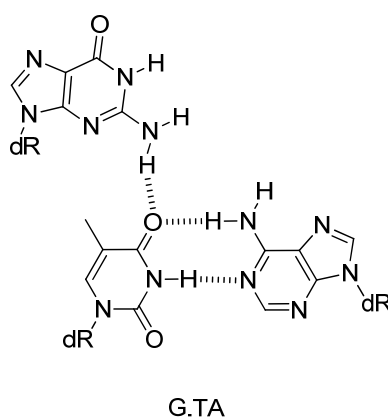


Fig. 1.49: The structure of the G·TA triplet.

Yoon *et al.*¹⁶⁷ first proposed the parallel triplet, T·CG (*fig. 1.50*). This triplet contains a single hydrogen bond between O-2 of thymine and the exocyclic amino group at C-4 of cytosine.¹⁷⁰ Up to three consecutive T·CG triplets can be accommodated within a DNA triplex, only if a triplex binding ligand is present, otherwise no footprint is evident.¹⁶⁸ To minimise backbone distortions, the sugar conformation of thymidine adopts a C^{3'}-*endo* pucker. The T·CG triplet can also be formed within a purine motif antiparallel triplex; this is due to the rotatable nature of thymine. The parallel and antiparallel forms of the T·CG triplet (*fig. 1.50*) are nearly superimposable; both possess *anti* conformation of the glycosidic bond, C^{3'}-*endo* sugar pucker and a slight 5' tilt by the thymine residue to maintain the π -stacking interaction along the helical axis.¹⁶⁵

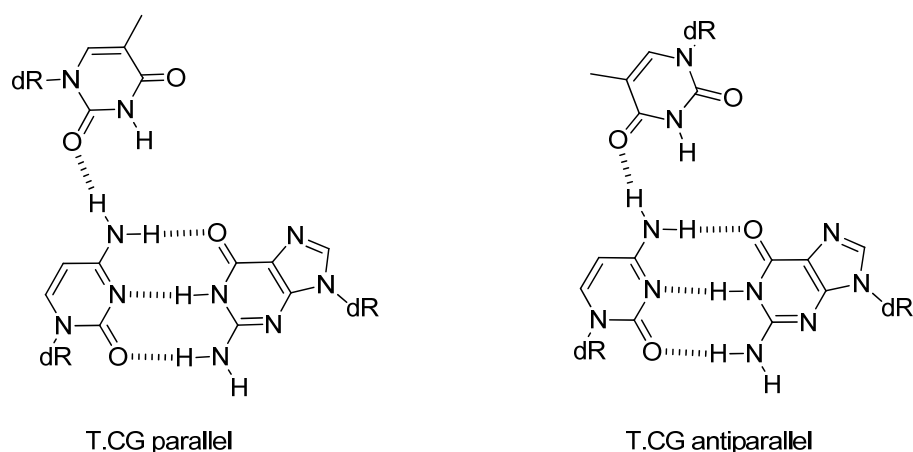


Fig. 1.50: The structures of the T·CG parallel and antiparallel triplets.

Abasic Linkers

The abasic linker, 1,2-dideoxy-D-ribose (ϕ) has been incorporated into TFOs to study its binding affinity and triplet stabilities when bound to the base pairs A·T, G·C, T·A and C·G.¹⁷¹ The abasic linker is not involved in base pairing or base stacking; it is simply present within the TFO to accommodate pyrimidine interruptions in the homopurine target strand. Within a pyrimidine third strand, the triplets ϕ ·AT, ϕ ·GC, ϕ ·TA and ϕ ·CG are significantly less stable than the triplets, T·AT, C⁺·GC and G·TA. The decrease in binding produced by an abasic site is similar to that observed with imperfectly matched natural base triplets. There exists a preference for T·A and C·G over A·T and G·C base pairs across the abasic site. Abasic linkers generally destabilise a triplex due to loss of base stacking and hydrogen bonding, hence they are unlikely to be used as a null position for the recognition of mixed sequence DNA sites by pyrimidine oligonucleotides.¹⁷¹

Universal Bases

A universal base analogue forms base pairs with each of the natural DNA/RNA bases, with little discrimination between them. Most of these analogues stabilise the DNA complex by stacking interactions; they are non-hydrogen bonding, hydrophobic, aromatic bases.¹⁷² This approach unfortunately has not yielded stable triple helical

structures; a loss of specificity is encountered at the universal base position as any base can be tolerated here.

Nucleotide Analogues for Recognising Pyrimidine Interruptions: Targeting CG

It has been found that in the antiparallel motif, only thymine shows substantial binding to C·G inversions¹⁷³ as mentioned above. O-4 of thymine is a major participant in the T·CG triplet (*fig. 1.50*). This was supported by the observation that a TFO containing pyridin-4-one deoxyribonucleoside (*fig. 1.51*) binds to the C·G inversion, while a TFO containing pyridine-2-one deoxyribonucleoside (*fig. 1.51*) does not. Binding of pyridine-4-one to C·G inversions is substantially weaker than T (by ~100-fold), hence, other interactions, such as hydrophobic and/or stacking effects are additionally involved. The importance of the O-4 group provided a starting point in the design of other T analogues with increased affinity and specificity for the recognition of C·G inversions.

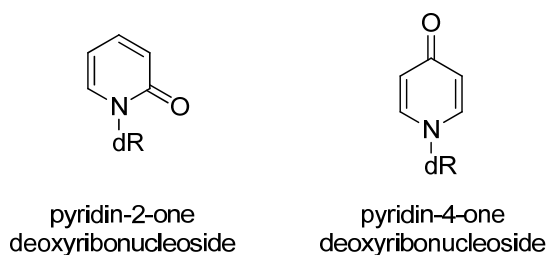


Fig. 1.51: The structures of the nucleosides pyrdin-2-one and pyridine-4-one.

It has been reported that the nucleobase 5-methyl-pyrimidine-2-one, ⁴H_T (*fig. 1.52*) can recognise cytosine of the C·G base pair in the parallel motif as efficiently as T, without the compromise in selectivity that is inherent to thymine.¹⁷⁴ The 4-carbonyl and 3-NH groups of T are used in the recognition of adenine, by omitting these groups, the selectivity for C·G could be increased alongside abolishing the recognition of A. Although selectivity was improved, the affinity was at a lower level compared to a conventional pu:pu:py triplet. The affinity of third strand binding is tuneable by using third strand bases that show enhanced binding to the purine bases (2-aminopyridine for recognition of G) and/or by using triplex specific intercalators. Favourable electrostatic interactions between the 2-carbonyl group of ⁴H_T and the 2-CH group at the 5-position

of cytosine may also contribute towards the stability of the $^4\text{H}\cdot\text{T}\cdot\text{CG}$ triplet; this unconventional hydrogen bond has been mentioned previously in assisting in the stability of the $^7\text{H}\cdot\text{GC}$ triplet.¹⁶⁴

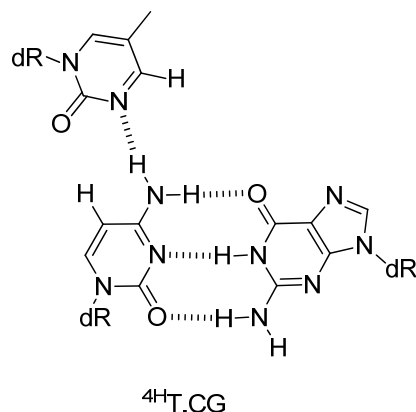


Fig. 1.52: The structure of the $^4\text{H}\cdot\text{T}\cdot\text{CG}$ triplet.¹⁷⁴

In order to increase the strength of the $^4\text{H}\cdot\text{T}\cdot\text{CG}$ triplet, Buchini *et al.*¹⁷⁵ modified the sugar with a 2'-*O*-aminoethyl side chain. This led to an increase in thermal stability of the triplex within which it was contained, of about 1.5 °C per modification. At pH 7.0, the stability of the 2'-modified triplex was higher than that of the triplex formed with the unmodified $^4\text{H}\cdot\text{T}$. Duplexes containing up to three C·G interruptions were able to be targeted by the corresponding TFOs containing the 2'-modified $^4\text{H}\cdot\text{T}$ base analogue at pH 6.5, however, with significantly reduced stability compared to the triplexes formed with the duplexes containing single C·G inversions. The formation of a triplex with five C·G inversions within the duplex was not observed, even by decreasing the pH to 5.5. However, on modifying the % pyrimidine content to 33 % (five C·G inversions) in the target duplex and using a TFO containing all 2'-*O*-aminoethyl ribonucleoside residues, Buchini *et al.*¹⁷⁶ were able to demonstrate recognition of a 15-mer DNA target duplex with high selectivity and a third strand affinity which is equal to, or greater than that of the two strands of the target duplex for one another.

6-(3-aminopropyl)-7-methyl-3H-pyrrolo[2,3-*d*]pyrimidin-2(7H)-one ($^{\text{A}}\text{PP}$), a derivative of $^4\text{H}\cdot\text{T}$ has also been reported^{112,177} to be selective for C·G interruptions in the parallel motif, and is ~ 4 °C more stable than T at pH 6.0. The $^{\text{A}}\text{PP}\cdot\text{CG}$ triplet (*fig. 1.53*) retains the H-bonding pattern of $^4\text{H}\cdot\text{T}\cdot\text{CG}$. The nucleobase introduces additional aromaticity,

and the protonated primary propylamino group can potentially provide extra stability by hydrogen bonding with N-7 or O-6 of guanine, or by participating in electrostatic interactions with the anionic phosphate groups.

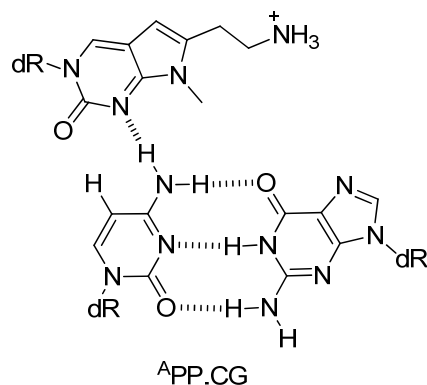


Fig. 1.53: The structure of the ^APPP·CG triplet.

Chen *et al.*¹⁷⁸ reported the use of the modified nucleobase, 2-aminopyrimidine (d2APm) in TFOs for the selective recognition of C·G inversions (*fig. 1.54*) in the parallel motif. d2APm has a pK_a of 3.3, hence providing an unprotonated nitrogen at all physiological pH which acts as the hydrogen bond acceptor site for the recognition of cytosine bases. At physiological pH, TFOs containing d2APm give higher triplex T_m values than the corresponding triplex formed with dC containing TFOs; ΔT_m values between +3 and +4 °C have been observed.

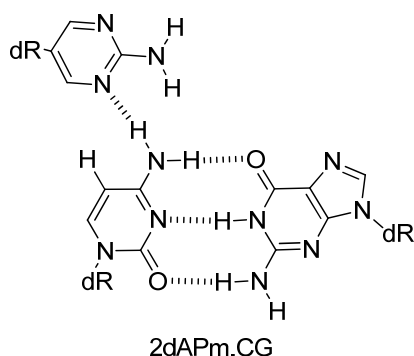


Fig. 1.54: The structure of the 2dAPm·CG triplet.

Obika *et al.*¹⁷⁹ reported the efficient and selective recognition of the C·G base pair by a bridged nucleic acid consisting of the 2-pyridone nucleobase, P^B (*fig. 1.55*). As a result

of the lack of a 3-N atom and a 4-carbonyl or amino group which are crucial for hydrogen bonding with other base pairs, P^B has reasonable selectivity for the C·G inversion. The 2',4'-bridged sugar modification produces a great level of stabilisation in comparison to the deoxy derivative; ΔT_m of +9 °C per modification.

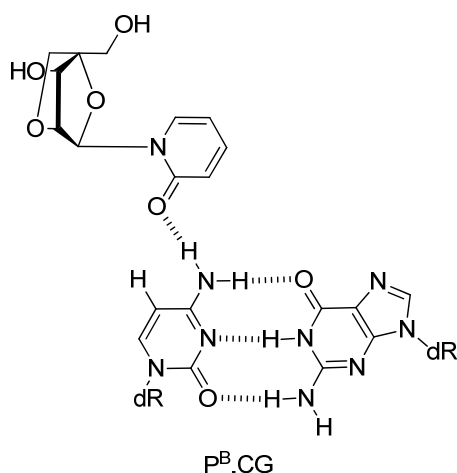


Fig. 1.55: The structure of the triplet P^B ·CG.

Nucleotide Analogues for Recognising Pyrimidine Interruptions: Targeting TA

The presence of the 5-methyl group in thymine presents steric limitations when developing analogues to interact with this base. Eldrup *et al.*¹⁸⁰ developed the PNA analogue, 3-oxo-2,3-dihydropyridazine (E) (*fig. 1.56*) which is connected to the PNA backbone *via* a β -alanine linker from the 6-position. This linker is long enough to potentially circumvent the 5-methyl group of thymine as was indicated by computer model building studies. Also, the lack of a hydrogen atom on N-1 greatly reduces any steric interference with this group. In comparison with guanine as an interacting base for the T·A base pair, the results indicated that the E-base binds stronger than guanine to thymine, with ΔT_m of +5 °C per modification. The E-base selectively binds to thymine over cytosine with ΔT_m of +7 °C per modification. However, in thermal stability studies conducted to compare uracil in the target in place of thymine, the E-base PNAs were shown to bind somewhat stronger to the uracil containing targets (ΔT_m of \sim +3 °C per modification). This indicated that the design of the E base does not completely solve the steric clash problem associated when a nucleobase in a TFO binds to thymine in the target strand.

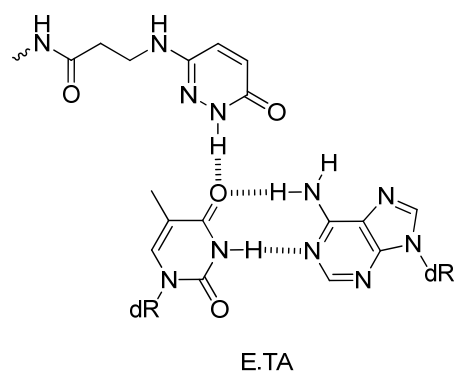


Fig. 1.56: The structure of the E·TA triplet.

It is thought that the reduced binding of the E-base to thymine compared to uracil was due to excessive flexibility of the backbone-nucleobase linker and therefore causing steric clash with the 5-methyl group of thymine. Olsen *et al.*^{181,182} addressed the flexibility issue by synthesising a new analogue, E^{ag} (*fig. 1.57*) which contains a double bond. The double bond was introduced to restrict the flexibility in the linker and thereby reduce the loss of entropy associated when E incorporated TFOs undergo triplex formation. The double bond was additionally thought to yield enthalpic stabilisation through stacking interactions with neighbouring base pairs. However, the thermal stability results demonstrated that the overall triplex formed with E^{ag} base is less stable than when using the original E-base; ΔT_m of $-1.5\text{ }^{\circ}\text{C}$ per modification. The recognition of uracil was different however; the E-base preferred binding to U over T, while the new base E^{ag} showed a minor preference to T compared to U, with ΔT_m of $+0.5\text{ }^{\circ}\text{C}$ per modification which suggested absence of significant steric clash between the E^{ag}-base and the 5-methyl group of thymine. In a comparison of the thermodynamic parameters of the two different bases, there was no significant difference; the entropy loss in particular, was very similar ($\Delta S^{\circ} = -2037\text{ KJmol}^{-1}$ for E^{ag} vs $\Delta S^{\circ} = -2058$ for E). Therefore, introduction of the more rigid E^{ag} base did not result in an entropic advantage.

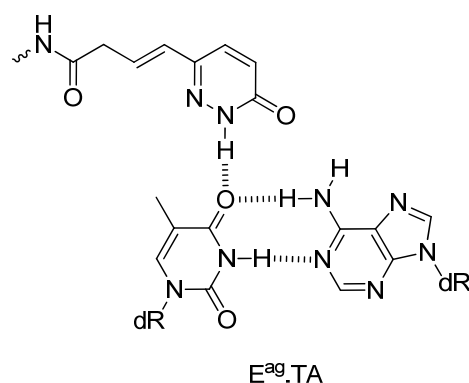


Fig. 1.57: The structure of the $E^{ag}.TA$ triplet.

Nucleotide Analogues for Recognising Both Partners of the Base Pair: CG Recognition

The non-natural deoxyribonucleoside 4-(3-benzamidophenyl)imidazole (D_3) (fig. 1.58) developed and reported by Griffin *et al.*¹⁸³ was designed to recognise C·G or T·A Watson-Crick base pairs, by sterically matching their edges, selectively within the parallel triplex motif. It was anticipated that the N-3 within the imidazole moiety, would form a single hydrogen bond to the 4-amino group of cytosine, whilst the presence of two aromatic rings positioned in the major groove would enable additional stacking interactions. The two aromatic rings were connected by a single bond providing a rotational degree of freedom, which allowed the non-natural base to adopt a favourable geometry for interaction with pyrimidine·purine base pairs. It was found that D_3 recognises both bases within the pyrimidine·purine base pair with preference in comparison to the purine·pyrimidine base pairs, within the parallel triplex motif. Neighbouring base triplets influence the stabilities of D_3 ·TA and D_3 ·CG by nearest neighbour interactions; the triplets are most stable when flanked by T·AT triplets and less stable when C^+ ·GC triplet is on the 3' side, hence introducing a limitation in that there is a need for consideration of the sequence composition of target sites. Another limitation with this nucleoside is that D_3 is not selective over either the T·A or C·G base pair.

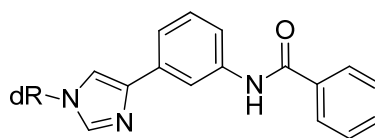


Fig. 1.58: The structure of the D₃ non-natural deoxyribonucleoside.¹⁸³

In fact, it was reported by Koshlap *et al.*¹⁸⁴ that D₃ binds to the T·A and C·G base pairs *via* intercalation at YpA; the binding mode of D₃ is unique in that it skips a potential base pair in order to intercalate.

Other analogues that share the sequence-selective intercalation as a common binding mechanism to C·G and T·A base pairs are **L1** and **L2** (*fig. 1.59*) which have been reported by Lehman *et al.*¹⁸⁵

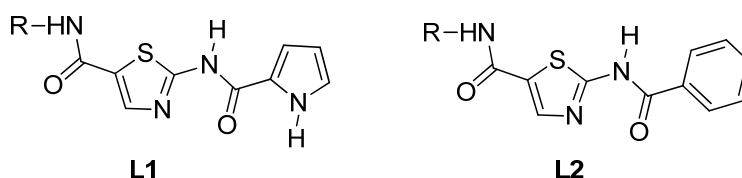


Fig. 1.59: The structures of **L1** and **L2**.

Huang *et al.*¹⁸⁶ reported the recognition of both bases of the C·G inversion, by N⁴-modified cytosine analogues, in triplex formation studies. The side chains of N⁴-(3-acetamidopropyl)cytosine, and N⁴-(3-aminopropyl)cytosine interact in a specific manner with the C·G base pair of a duplex. For both nucleosides, the 3-acetamide/3-amino hydrogens, respectively, are able to hydrogen bond to the O-6 carbonyl of guanine of the target C·G base pair, as a result of sufficient side chain length. Triplex formation of N⁴-(3-acetamidopropyl)cytosine is selective for the C·G base pair ($T_m = 20$ °C); this modified nucleoside has no affinity for A·T or T·A ($T_m < 0$ °C); some affinity is observed when a G·C base pair was present ($T_m = 8$ °C). UV melting studies demonstrated that the triplet composed of this base with C·G is more stable than C·CG but less stable than the canonical triplets.¹⁸⁶

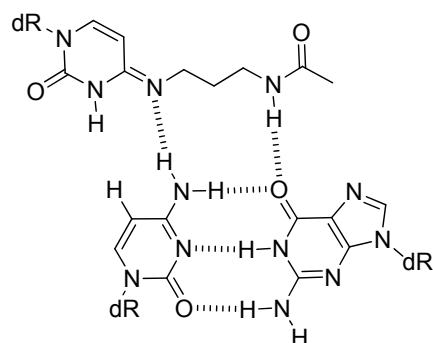


Fig. 1.60: The structure of the triplet formed between N^4 -(3-acetamidopropyl)cytosine and the C·G base pair.¹⁸⁶

The N^4 -(6-aminopyridinyl) derivative (*fig. 1.61*) has also been reported to form stable triplexes with DNA duplexes containing C·G base pair inversions.^{187,188} Hydrophobic interactions between the pyridinyl ring of N^4 -(6-aminopyridinyl)cytosine and neighbouring bases may contribute to triplex formation.

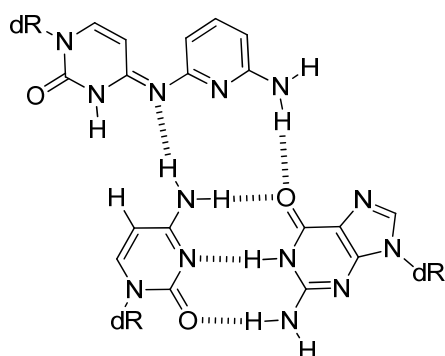


Fig. 1.61: The structure of the triplet formed between N^4 -(6-aminopyridinyl)cytosine and the C·G base pair.^{187,188}

Nucleotide Analogues for Recognising Both Partners of the Base Pair: TA Recognition

A novel nucleobase, S (*fig. 1.62*), was reported by Guianvarc'h *et al.*¹⁸⁹ to effectively circumvent a T·A base pair interruption within a purine·pyrimidine duplex when incorporated into a pyrimidine-motif TFO. This nucleoside was achieved through an attempt to achieve better triplex stabilisation than the D_3 monomer at pyrimidine·purine inversions. The S nucleobase consists of two unfused aromatic rings which are linked to 2'-deoxyribose by an acetamide motif. Three hydrogen bonds are present in the S·TA

triplet (*fig. 1.62*). Thermal denaturation experiments indicated that the T_m value of the triplex containing the S·TA triplet ($T_m = 50\text{ }^\circ\text{C}$) was very close to those of perfect triplexes without any interruptions of the homopurine·homopyrimidine sequences (T_m (T·AT) = $51\text{ }^\circ\text{C}$).¹⁸⁹ There does however exist steric hindrance between the 5-methyl group of T in the oligopurine-rich target strand and the deoxyribose moiety of S, as was demonstrated by the increased T_m /thermal stability obtained by the triplex formed when T was replaced by U, featuring the S·UA triplet. Triplexes containing S are most stable when the S·TA triplet is flanked by two T·AT triplets, and the stability of the triplexes decrease as the number of adjacent C⁺·GC triplets around the S·TA triplet is increased. In comparison to the best natural base triplet, G·TA, the S·TA triplet provides better triplex stabilisation.¹⁹⁰

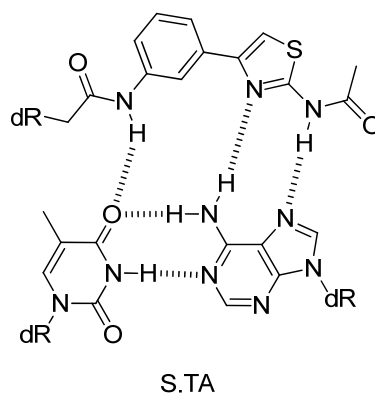


Fig. 1.62: The structure of the S·TA triplet.

B^t , a nucleobase analogue featuring a thiazolyl-benzimidazole system, is a derivative of the S base. It was anticipated by Guianvarc'h *et al.*¹⁹¹ that this modified nucleobase would form three selective hydrogen bonds with a T·A inversion (*fig. 1.63*). B^t has an extended aromaticity of the benzimidazole ring, this design was based on the rationale that it would achieve greater rigidity of the base, which would hence better stabilise triplex formation. Hence, it was expected that B^t would exhibit more efficient triplex stabilising properties than S, as a result of the favourable contributing entropic factors. However, in contrast to theoretical prediction, thermal denaturation studies gave a T_m value for the triplex containing a B^t ·TA triplet ($T_m = 43\text{ }^\circ\text{C}$) to be lower than that for S·TA, and very close to that obtained for a triplex containing a G·TA triplet ($T_m = 45\text{ }^\circ\text{C}$). From these results, it was concluded that B^t did not involve three hydrogen bonds as was first thought, this was based on the results obtained for triplexes

containing the G·TA triplet that contains only one hydrogen bond and a higher T_m . The B^t ·TA triplet is thought to involve non-sequence specific interactions such as third strand stacking/intercalation. In addition, B^t induced only a moderate discrimination between the T·A and C·G base pairs, demonstrating non-specificity of the nucleobase, further indicating interactions involved with B^t to be that of intercalation.

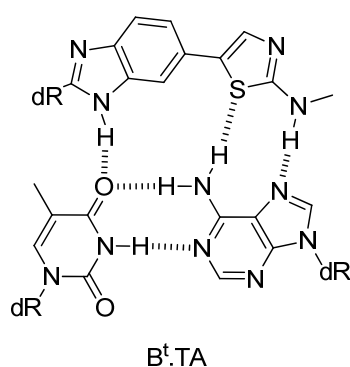


Fig. 1.63: The structure of the B^t ·TA triplet.

Wang *et al.*¹⁹² prepared the 2'-aminoethoxy derivative of the S nucleoside (S_{AE}) (fig. 1.64) and reported its incorporation into triplex forming oligonucleotides for recognition of T·A interruptions within a target oligopurine tract.^{112,192} Biophysical studies indicated that S_{AE} had a greater affinity for a single T·A interruption than either G, or S. In the presence of eight C^+ ·GC triplets, stable triplexes were formed at pH 6.0 at an 18-mer target site containing two T·A interruptions. In addition to producing stable triplexes at T·A interruptions, S_{AE} and S also produce stable triplexes with other base pairs, in particular C·G, although the selectivity for T·A improves with increased pH.

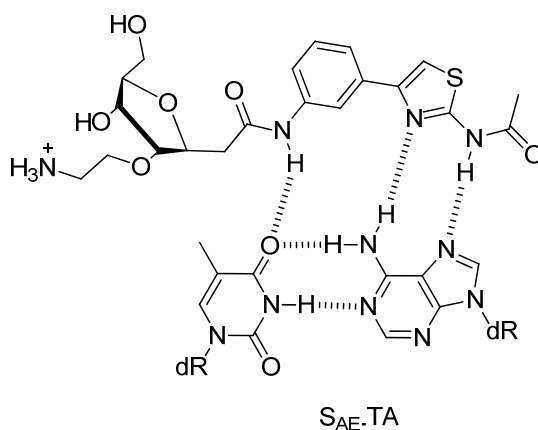


Fig. 1.64: The structure of the $S_{AE} \cdot TA$ triplet.

Craynest *et al.*¹⁹³ suggested the recognition of the T·A inversion by extended guanine analogues, **N1**, **N2**, **N3** (fig. 1.65), derived from aminobenzimidazole (recognition of the thymine base) and thymine or 5-substituted uracil (for recognition of the adenine base) by three hydrogen bonds.

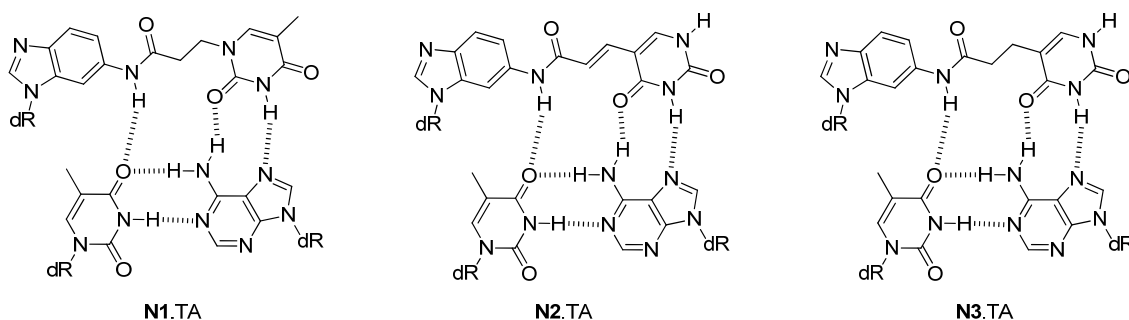


Fig. 1.65: The structures of the **N1**·TA triplet, **N2**·TA triplet and the **N3**·TA triplet.

Another important factor to consider when designing antigene oligonucleotides is the accessibility of the target sequence within the chromatin structure of the cell nucleus. TFOs can encounter many obstacles en route to their genomic target sequence, these include; the nuclear chromatin structure/nucleosomal organization; as well as the potential trapping of the TFO by nuclear proteins which results in a decrease in the concentration of TFO that reaches the target.¹⁹⁴

1.7 METHODS TO STUDY THE BIOPHYSICAL PROPERTIES OF NUCLEIC ACIDS

1.7.1 Ultra-Violet (UV) Melting^{195,196}

UV-melting analysis is a useful technique to study the biophysical properties of nucleic acids, and hence understand their role in biological function.

According to the molecular structure of a substance, light of specific wavelengths can be selectively absorbed. The wavelength of light (λ) absorbed is inversely related to the energy difference (see eqn. 1) between two allowed states of the valency electrons of that particular molecule.

$$E = hc/\lambda$$

Equation 1: where c denotes the speed of light, h is Planck's constant, and λ is the wavelength of light..

The incident photon of specific energy promotes an electron from the ground state to an excited state. The excited electrons eventually return to the ground state, by loss of energy *via* heat radiation.

Spectrophotometry is a method used to quantitatively measure the absorbing species (chromophores), and is based on the Beer-Lambert's Law. Lambert's law states that '*the fraction of light absorbed by a transparent medium (I/I_0) is independent of the incident light intensity (which is dependent on the distance the light has had to travel), and each successive layer of the medium absorbs an equal fraction of the light passing through it*'.¹⁹⁵ This is expressed mathematically as follows:

$$I/I_0 = 1/10^{kl}$$

rearranges to

$$I_0/I = 10^{kl}$$

$$\log_{10}(I_0/I) = kl$$

Where I_0 is the intensity of the incident light, I is the intensity of transmitted light, l is the length of the light-path in the spectrophotometer cuvette, and k is a constant for the medium, which is deciphered by Beer's law.

Beer's Law claims that '*the amount of light absorbed is proportional to the number of molecules of the chromophore through which the light passes*'.¹⁹⁵ So, the constant k is proportional to the concentration (c) of the chromophore: $k = \epsilon c$, where ϵ is the molar absorption coefficient, a property of the chromophore itself, it is numerically equal to the absorption of a molar (1 mol L^{-1}) solution in a 1 cm light path. Hence, the expression of the combined Beer-Lambert law:

$$\log_{10} (I_0/I) = \epsilon cl$$
$$\text{or } A = \epsilon cl$$

Where the term $\log_{10} (I_0/I)$ is called the absorbance (A).

The purine and pyrimidine nucleobase components of nucleic acids are chromophores; they absorb light energy of wavelength between 240-280 nm at neutral pH, with an absorption maximum (Abs_{Max}) at around 260 nm. This corresponds to UV light. The absorption coefficient of each DNA nucleotide residue is constant; hence the absorption at 260 nm can be used to determine the nucleic acid concentration. The absorption of a native nucleic acid molecule (oligonucleotide strand) cannot however, be expressed as a sum of absorbances of all the component nucleotides. It is actually less than the sum of the absorptions of its constituent nucleotides. This is known as the *hypochromism* of nucleic acids,^{197,198} and arises from the weak interactions (π -stacking) between the neighbouring residues.

When nucleic acids are exposed to high temperatures, extreme pH, low ionic strength, or high concentration of urea, they can become denatured. The denaturation process involves unwinding of the double helix of the native DNA, the two strands separate as random coils, and weak interactions between neighbouring nucleotides are diminished. This results in an increase in the molar absorption coefficient, because single stranded DNA absorbs UV light more effectively than double stranded DNA, and thus the spectra of nucleic acids are associated with a sharp increase in the absorption of UV light.¹⁹⁷ Hence, the melting temperature of DNA can be determined by measuring the extent of the hyperchromic effect as a function of temperature.

Stacked bases have a smaller absorption per base than unstacked bases; this is called *hyperchromicity* which is defined as:

$$\% \text{ hyperchromicity} = 100 \frac{(A_{\text{denatured}} - A_{\text{native}})}{A_{\text{native}}}$$

where $A_{\text{denatured}}$ and A_{native} are the absorbances at high and low temperature, respectively. The absorbance versus temperature profile is commonly referred to as a UV absorbance melting curve (*fig. 1.66*).

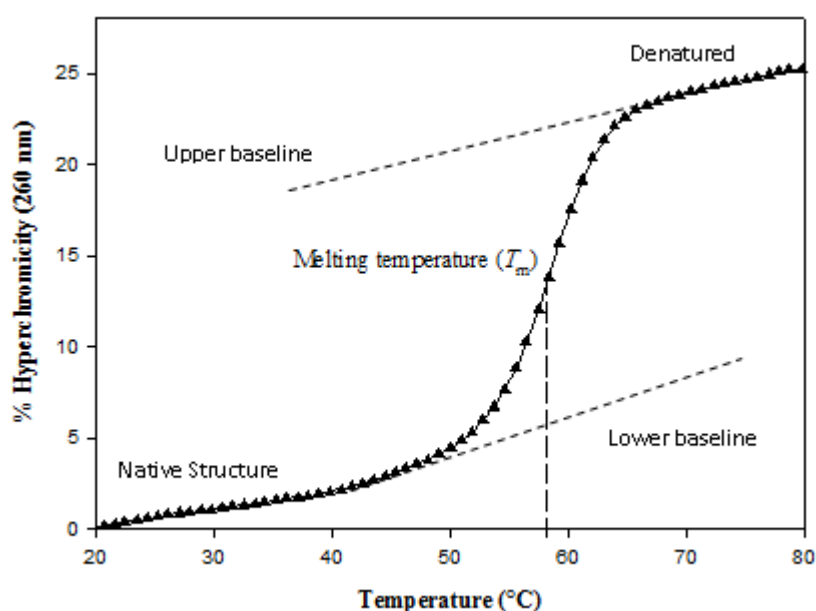


Fig. 1.66: Typical experimental UV melting profile.

As the temperature increases, the UV absorbance/% hyperchromicity increases, this is due to an increase in the ratio of molecules in the single stranded version versus native state. The melting temperature, T_m , is defined as the temperature at which half of the strands are in the native state and half are in the ‘random coil’ state. UV absorbance is a highly sensitive technique due to the high molar absorption of the bases.¹⁹⁵

The formation of a DNA triple helix between a triplex-forming oligonucleotide (TFO) and double stranded DNA (dsDNA) is a very important process that can be potentially exploited for artificial regulation of gene expression. Analysis of DNA-triplex

formation *in vitro* is the first step for any study regarding practical applications of triple helices.

Triplex DNA thermal denaturation (*fig. 1.67*) is based on the hyperchromic effect upon the association of the third strand to its target DNA.

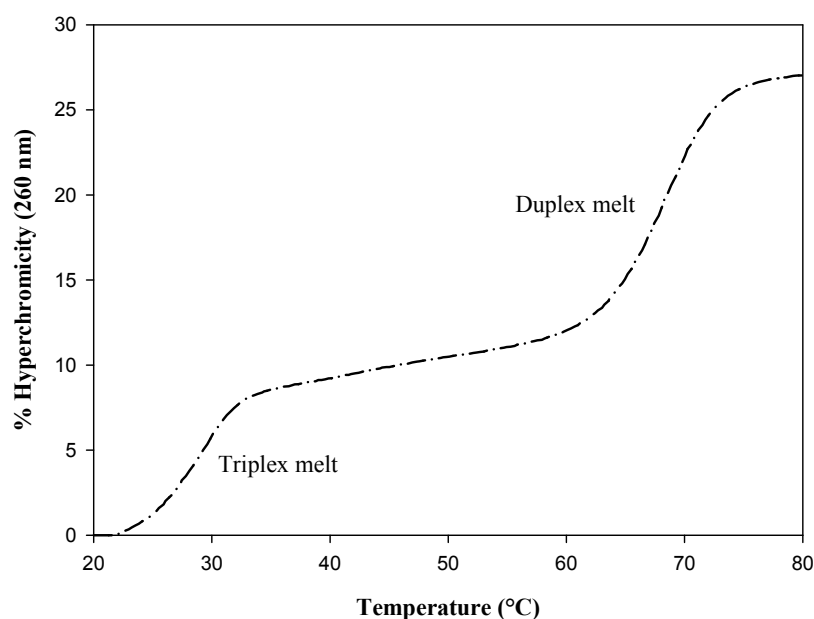


Fig. 1.67: A typical UV melting curve showing the biphasic transition of a triplex melt.

Here, the melting temperature (T_m) is the temperature where half of the triplex has dissociated into duplex plus third strand. The extinction coefficient of the triple helix is lower than that of the dissociated complex, so UV absorbance is increased upon heating (dissociation) of the triplex. This increase is cooperative and has a sharp transition that permits detection of the melting temperature with quite high precision.¹⁹⁶

Both intrinsic factors such as length and sequence composition, and extrinsic factors such as ionic, pH and temperature, influence the stability of DNA triplexes and therefore the melting temperature (T_m). Hence these factors need to be taken account of when designing sequences and carrying out UV-melting analysis.

Sequences with runs of 3 or more guanines should be avoided, since they tend to form G-quartet structures. It is also useful that sequences be designed to have T_m values

between 20 °C and 75 °C, to allow upper and lower baselines to be adequately defined. With regards to hairpins, to minimize bimolecular self-complementary internal loop formation, it has been recommended that a sodium concentration of below 0.1 M is used. Self-complementary sequences have a high propensity to form hairpin structures that compete with the bimolecular reaction; hence these should also be avoided. If self-complementary sequences are used, it is recommended that measurements be carried out in high salt conditions (e.g 1M NaCl) and to design the middle of the duplex to be G-C rich and the ends of the duplex to be A-T rich. The very ends of the duplexes, however, should be G-C whenever possible to minimize ‘end-fraying’ artefacts.¹⁹⁵ The ideal length of the target sequence is usually between 20 and 45 bp. It has been shown that sequences adjacent to a triple helix can influence its stability, hence to be able to fully analyse how well a TFO binds to its target, the target should ideally be put in its natural context. If the triple helix is very stable, a longer target will be preferred in order to clearly differentiate between the melting of the triplex and the dissociation of the duplex. The stability of a hairpin duplex can be increased substantially, by the covalent linkage of both the strands, the melting temperature is normally achieved at temperatures above 90 °C.¹⁹⁶ Choice of buffer with an appropriate pK_a to give sufficient buffer capacity at the desired pH is important.¹⁹⁵ It is also important that samples are pure (typically > 95 % purity), free of dust and degassed, this enables the accurate determination of nucleic acid thermodynamics.¹⁹⁵

UV-melting analysis is a widely used method for the determination of triple-helix stability, especially for the study of chemical modifications in the third strand.¹⁹⁶

1.7.2 Footprinting¹⁹⁹

Footprinting is a method used to assess the sequence selectivity of DNA-binding ligands, estimate their binding strengths (quantitative footprinting) as well as their association and dissociation rate constants for slow binding reactions. The ligands include polyamides, triplex-forming oligonucleotides and minor groove binding ligands.

Footprinting utilises cleavage agents such as DNase I or hydroxyl radicals to digest double stranded DNA; these agents can modify and/or cleave the phosphodiester backbone.²⁰⁰ When a ligand is bound at specific binding sites within a DNA fragment,

digestion is locally inhibited, because the ligand protects the section of the backbone to which it is bound, from the cleavage agents.¹⁹⁹ Footprinting is therefore a protection assay (see *fig. 1.68*).

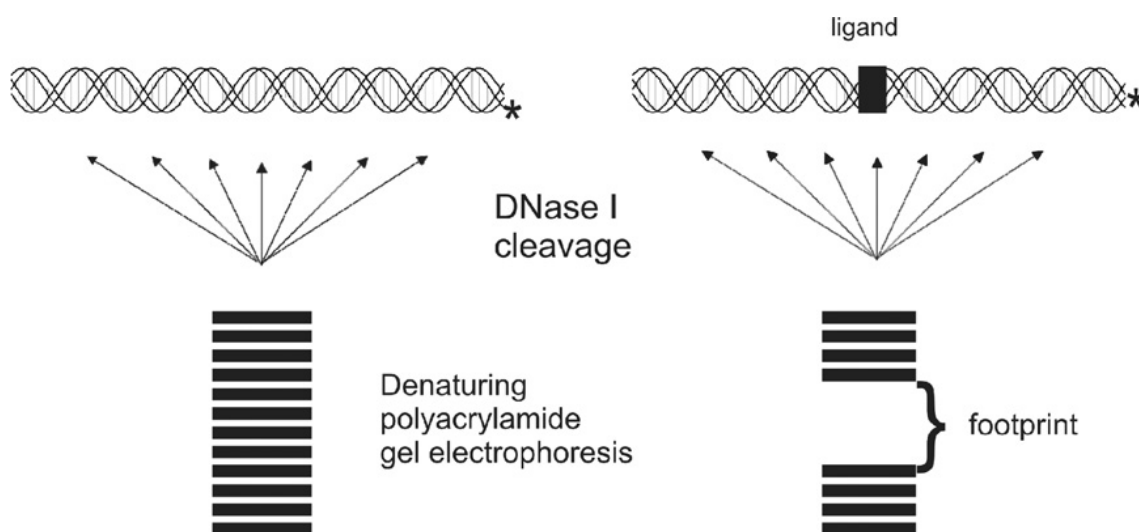


Fig. 1.68: Schematic representation of the footprinting experiment. The asterisk represents where the DNA fragment is labelled. Here, DNase I is the cleavage agent used to digest the DNA under conditions of single-hit kinetics. The ligand protects the DNA fragment from cleavage at its binding site, this is evident as a ‘footprint’ when the products of digestion are resolved on a denaturing polyacrylamide gel.¹⁹⁹

A double-stranded DNA fragment is radiolabelled with either ^{32}P or ^{33}P , at either the 3’-end or 5’-end of one strand and is cut by a chemical or enzymatic cleavage agent, on average each DNA molecule is cleaved only once (i.e. single-hit kinetics). A random distribution of products is produced when the cleavage agent does not possess any sequence selectivity, and this can be resolved on a denaturing polyacrylamide gel. In the presence of a sequence selective ligand, the cleavage agents cannot cause damage to that particular region of the DNA as it is protected. The cleaved DNA strand fragments are resolved on the gel according to their length, with the longest fragments residing at the top. Sequences that include successive bases from the target site will be missing from the reaction and will therefore be evident as a gap (“footprint”) in the gel.¹⁹⁹

Footprinting can be used to determine the exact binding sites for a ligand, by running control and ligand-treated digestion alongside suitable markers (e.g. a Maxam–Gilbert chemical sequencing reaction). In addition, a ligand’s binding affinity can be estimated by performing the footprinting reaction over a range of ligand concentrations.¹⁹⁹

DNase I and hydroxyl radicals are the most widely used cleavage reagents in footprinting studies. DNase I acts by introducing single strand nicks into the phosphodiester backbone, cleaving the O3'–P bond, it is a double-strand-specific endonuclease. Cleavage by DNaseI is activated by divalent cations; Mg²⁺/Ca²⁺ are the usual ions of choice (there is a 10-fold more efficiency with magnesium than calcium). The cleavage efficiency can be enhanced by manganese, and so it is usually included into the reaction buffers. EDTA is used to stop the reaction by chelating these ions. Hydroxyl radicals are highly reactive freely diffusible species, they are generated by the Fenton reaction between Fe²⁺ and H₂O₂ (Eq. 2).¹⁹⁹



They are able to cleave close to the bound ligand as a result of their small size, and so they more accurately estimate the position of the ligand binding site. The cleavage reaction is performed in the presence of EDTA, which chelates the iron as [Fe(EDTA)]²⁻, this is to prevent direct interaction between Fe²⁺ and the negatively charged DNA substrate. It is thought that the hydroxyl radicals attack at the C4' or C1' positions (via the minor groove) leading to subsequent base removal and strand scission.¹⁹⁹

Footprinting substrates are typically restriction fragments that have been obtained from appropriate plasmids; they are usually 50–200 base pairs long. It is essential that the fragment should contain the preferred binding sites for the ligand under investigation.¹⁹⁹

1.8 AIMS OF THIS PhD RESEARCH

The aims of this PhD research was to firstly develop an efficient, and novel synthetic route to the non-naturally occurring cytidine analogue, 3-methyl-2-amino-5-(2'-deoxy-β-D-ribofuranosyl)pyridine (^{Me}P). Although synthetic routes to this nucleoside have previously been reported^{158,201,202} (as discussed above and in *chapter 2*), these procedures are lengthy and have significantly low overall yields. The potential benefits of developing a shorter route with greater overall yield include extensive usage of this monomer in various TFOs for the study of triplexes in the pyrimidine motif. In particular, TFOs that are covalently attached to crosslinking photoreagents such as

psoralen (see *section 1.9*) can be used to target diseases that arise from point mutations, such as psoriasis (see *section 1.10*), for gene silencing therapeutics. The second aim of this research was to incorporate ^{Me}P and psoralen into TFOs that target part of the *gfp* (green fluorescent protein) coding sequence in transgenic *Caenorhabditis elegans* (*C. elegans*) (see *section 1.11*), and carry out *in vitro* biophysical analyses (UV melting, footprinting and gel retardation; see *chapter 3*). If these *in vitro* studies proved successful, *in vivo* studies would be carried out in the transgenic worms, to observe whether the expression of *gfp* had been inhibited. The third aim of this research was to incorporate ^{Me}P as cytidine equivalents into a TFO that target part of the *SLO-1 promoter* gene in *C. elegans*, and compare the affinity of this TFO with TFOs containing other cytidine analogues (5-methyl cytidine (^{Me}C), 3-methyl-2-amino-5-(2'-*O*-methyl-β-D-ribofuranosyl)pyridine (^{Me}P_{OMe}), and 5-methyl-(2'-*O*-aminoethyl-β-D-ribofuranosyl)cytosine (^{Me}C_{AE}), at physiological pH. This was to determine, which cytidine analogue out of these four, when incorporated into TFOs, gives the better triplex stability, higher binding affinity, and least pH dependency. The final aim of this research was to incorporate ^{Me}P and psoralen into a TFO that targeted part of the *UNC-22* coding gene within *C. elegans*, and study the *in vitro* biophysical properties of the generated triplex. If the results for these analyses proved successful, this TFO will be used *in vivo* in *C. elegans*, to study the effect (if any) this TFO has on this organism, intracellularly.

Overall, this research will contribute to the area of nucleic acid chemistry, and gene silencing therapeutics. Whether the results are positive, or negative, the information obtained will enable further research and the development of more robust techniques and study designs.

1.9 PSORALEN²⁰³

Psoralens are natural products found in plants. They are unique in their ability to freeze helical regions of DNA. The reaction of psoralen with DNA occurs by a two-step mechanism. The planar psoralen molecule first intercalates within a double helical region of DNA. Irradiation into an absorption band of the psoralen molecule controls its covalent addition to DNA. Covalent adducts form with the pyrimidine bases of nucleic acids, they are stable, but photoreversible. These covalent adducts can form at one or

both ends of the psoralen molecule; they are therefore bifunctional photoreagents. In using this photoreaction to form covalent crosslinks with base paired structures, psoralens can be used to stabilize DNA triplex structures.²⁰³

Psoralens are formed by the linear fusion of a furan ring with a coumarin; they are tricyclic compounds (*fig. 1.39 in section 1.6.4.1.4*).²⁰³

When psoralens undergo covalent photocycloaddition, they first intercalate into double-stranded DNA and then react at either the furan or coumarin end, primarily with thymidine, under incident light of wavelength of 320-400 nm. A covalent interstrand crosslink forms when both ends of a psoralen molecule bind to thymidines in opposite strands of a DNA helix.²⁰³

A conformational change in the DNA backbone is required in order to accommodate a diadduct; this is a result of the dihedral angle that exists in a psoralen monoadduct between the plane of a psoralen molecule and the plane of a pyrimidine base. Provided that the photoreaction between psoralen and DNA is carried out under anoxygenic conditions, there is no additional degradation of DNA. During the photochemistry reaction, if singlet oxygen is produced, it will degrade all biological molecules with which it comes in contact.²⁰³

The degree of psoralen photoreaction/the ratio of monoaddition: crosslinkage can be controlled by light dose, by either selecting suitable wavelengths to induce the chemistry, or by controlling the duration of light delivered. The photoreaction between psoralen and pyrimidines can be reversed using short wavelength ultraviolet light (254 nm). The monoadduct and crosslink are chemically stable under various chemical conditions, hence, the photochemistry can be carried out under a wide variety of conditions including broad ranges of temperature and ionic strength, in the presence or absence of divalent cations, and even in some organic solvents.²⁰³

Psoralens are diverse photoreagents; they can be used clinically as photochemotherapeutic agents for the treatment of psoriasis (*see section 1.10*), vitiligo, and other skin disorders. Psoralens can also be used to inactivate viruses and other pathogens, by irreversible nucleic acid specific crosslinking. Psoralen photochemistry

can also be used to inactivate viruses that cannot be otherwise inactivated for vaccine production.

Psoralens are ideal probes for nucleic acid structure wherever long wave ultraviolet light can be delivered, this is because they are able to penetrate most biological structures and are not highly toxic to cells in the absence of actinic light. DNA structure has been probed by psoralen crosslinking.²⁰³

1.10 PSORIASIS²⁰⁴⁻²⁰⁷

Psoriasis is a chronic condition, for which the cause is unknown; however, there appears to be a genetic predisposition.²⁰⁴

Psoriasis is principally a T-cell (type of immune cell) mediated disease; epidermal abnormalities are consequent upon the release of various pro-inflammatory T-cell derived cytokines.²⁰⁵

The skin is made up of two main layers: the outer (epidermis) and lower (dermis). Epidermal cells reproduce at the lowest part of the epidermis and take 28 days to travel to the surface. In psoriasis, they divide at a faster rate due to the release of the T-cell derived cytokines, and move up to the surface in about four days. Since the cells are not fully mature and keratinised, they stick to each other and build up on the skin surface, resulting in a scaly appearance. The capillary loops in the dermis become dilated and the skin red and inflamed; white blood cells leak from these dilated capillary loops into the skin.²⁰⁴

Moderate to severe psoriasis has been successfully treated with a combination of psoralen and long-wave ultraviolet radiation (PUVA). Oral psoralen is taken two hours before exposure to UV radiation.²⁰⁶ When the psoralen-treated skin is exposed to UVA light, the psoralen binds to thymidine in the DNA of the skin cells and slows DNA replication; it induces site-specific chromosomal breaks within the genome (*fig. 1.69*). The resultant DNA damage will stimulate the genomic repair system and thereby increases the frequency of homologous recombination. Eventually cell division slows in the epidermis.²⁰⁷

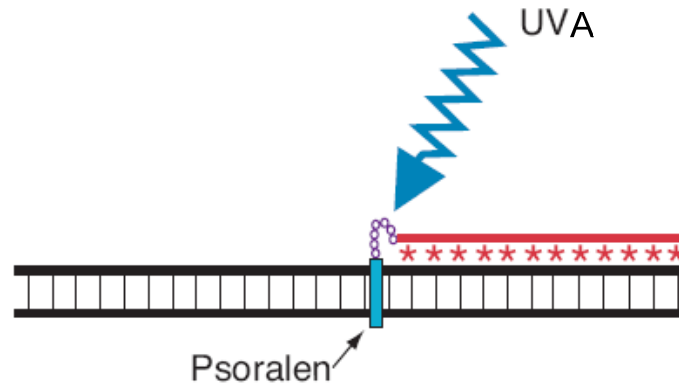


Fig. 1.69: TFO mediates site-specific delivery of the mutagen, psoralen (turquoise) linked to the oligonucleotide. The psoralen intercalates into the adjacent genomic DNA, and when activated by UVA irradiation, cross-links the three DNA strands, usually at adjacent thymines. Subsequent repair of the lesion via nucleotide excision repair results in base conversion.⁷⁵

1.10 CAENORHABDITIS ELEGANS (*C. elegans*)

The use of DNA triplexes for targeted modification of gene activity is poorly understood, and fundamental questions still need to be addressed.

C. elegans are an ideal organism for the study of gene regulation and function. *C. elegans* is a eukaryote, an organism whose cells contain complex structures inside the membranes. The genetic material is carried within the nucleus. *C. elegans* share their cellular and molecular structures (membrane bound organelles; DNA complexed into chromatin and organized into discrete chromosomes, etc.) and control pathways with higher organisms. They are multicellular organisms, which means that they go through a complex developmental process, including embryogenesis, morphogenesis, and growth to an adult. Thus, biological information that is learnt from *C. elegans* may be directly applicable to more complex organisms, such as humans.²⁰⁸ About 35 % of *C. elegans* genes have human homologs (a DNA or protein sequence that is similar to another DNA or protein sequence because the sequences have common ancestry). Many *C. elegans* genes can function similarly to mammalian genes. They have a fast and convenient life cycle. Embryogenesis occurs in approximately 12 hours, development to the adult stage occurs in 2.5 days, and the life span is 2-3 weeks.²⁰⁹ The development of

C. elegans is known in great detail because this tiny organism (1 mm in length) is transparent and the developmental pattern of all 959 of its somatic cells has been traced.

Surprisingly, the use of TFOs to modulate gene activity in *C. elegans* has not as yet been studied.

It has previously been demonstrated that dsRNA has the ability to act systematically within *C. elegans*, by using a ‘soaking’ method; dsRNA can get access into *C. elegans*, and cause gene interference simply by soaking the animals in solutions (1 to 5 µg/mL) of the dsRNA.²¹⁰ This gives confidence that TFOs can be designed to gain access to cells, following external application (soaking) to the whole animal. Although dsRNA and single stranded TFOs may use different uptake mechanisms, modifications can be introduced at the 2’-position to make these molecules more RNA-like, as well as providing stability against nucleases.

CHAPTER 2

Synthesis of the *C*-Nucleoside Cytosine analogue 3-methyl-2-aminopyridine deoxyriboside (^{Me}P)

2. SYNTHESIS OF THE C-NUCLEOSIDE CYTIDINE ANALOGUE 3-METHYL-2-AMINOPYRIDINE DEOXYRIBOSIDE (^{Me}P)

2.1 INTRODUCTION

The ability to target specific sequences of DNA through oligonucleotide-based triple-helix formation provides a potentially powerful tool for gene silencing.⁷⁵ However, within the pyrimidine motif, the requirement for the N3 protonation of the third strand cytosine (C^+) is one of the major limitations to using naturally existing bases in TFOs. Cytosine protonation requires low pH (due to the relatively low pK_a of deoxycytidine, $pK_a = 4.3$) to form the two Hoogsteen hydrogen bonds that are required for the formation of the $C^+.GC$ triplet (*fig. 2.7*). For gene silencing, TFOs must be active at physiological pH, at which the $C^+.GC$ triplet will not form.

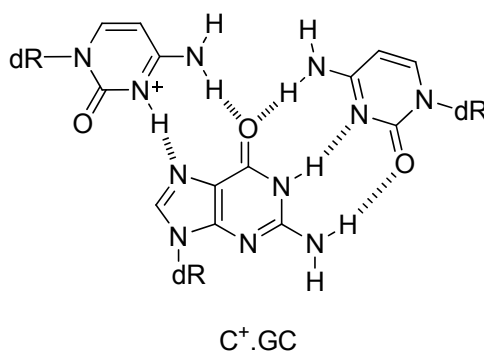


Fig. 2.1: Structure of the pyrimidine motif $C^+.GC$ triplet.

Several cytosine analogues have therefore been designed with the aim of forming Hoogsteen hydrogen bonds with guanines of GC base pairs at physiological pH. These have already been discussed in Chapter 1 (*section 1.6.4.2*). The cytosine analogue under investigation in this study is 2-amino-3-methyl-5-(2'-deoxy- β -D-ribofuranosyl)pyridine (^{Me}P) **1** (*fig. 2.2*), which is a C-nucleoside.

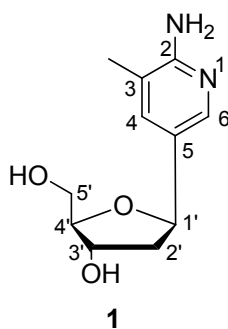


Fig. 2.2: The structure of the cytosine-analogue *C*-nucleoside, ^{Me}**P** (**1**).

Hildbrand *et al.*²¹¹ and Bates *et al.*²¹² have independently suggested the replacement of 2'-deoxycytidine, by the structurally closely related *C*-nucleoside, 2-amino-5-(2'-deoxy-β-D-ribofuranosyl)pyridine (**P**) (*fig. 2.3*) in TFOs.

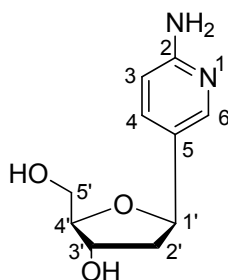


Fig. 2.3: The structure of the cytosine-analogue *C*-nucleoside, **P**.

For protonation to occur at higher pH, the basicity of the nucleobase needs to be increased. By removing the 2-oxo function, and replacing the N¹ (numbering taken from cytosine numbering system) in the pyrimidine ring of cytosine with a carbon, the p*K*_a of **P** is increased to 5.93,¹⁵⁶ which is closer to physiological pH, whereas that for cytosine is 4.3. **P** is structurally analogous to dC; it can still form two Hoogsteen hydrogen bonds to guanine bases in duplex DNA.

^{Me}**P** was derived from **P**; this idea was based on the fact that ^{Me}**C** showed enhanced triplex stability results when incorporated into oligonucleotides as deoxycytidine (dC) equivalents. Although ^{Me}**C** has a slightly higher p*K*_a (4.5) than cytosine, the enhanced triplex stability is thought to result from favourable hydrophobic effects generated by the formation of a spine of methyl groups in the major groove. This spine of methyl groups releases water molecules from the major groove, resulting in a favourable increase in the entropy of the surrounding solvent.^{131,158} ^{Me}**P** is ~ 0.4 p*K*_a units more

basic than **P**. In the triplex melting studies carried out by Hildbrand *et al.*¹⁵⁸, they observed no difference in triplex stability between triplexes formed with oligos containing **P** and those containing ^{Me}**P**. Hence, this result suggests that the binding of cytosine to GC base pairs is best improved by enhancing the degree of N-3 protonation (by increasing the pK_a), rather than by attaching a hydrophobic methyl group to the C-5 position.

^{Me}**P** is able to bind to the GC base pair by the formation of two hydrogen bonds (*fig. 2.4*), without perturbing the hydrogen-bonding pattern in neighbouring triplets that are relevant for triple-helix formation.

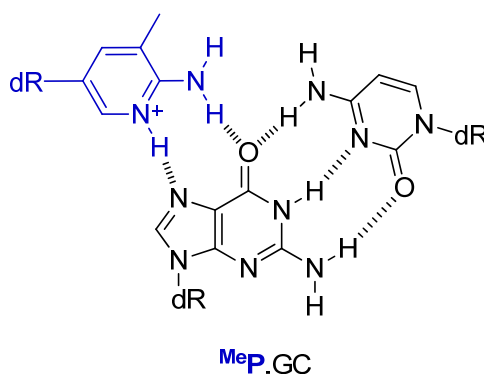


Fig. 2.4: Structure of the pyrimidine motif ^{Me}**P**.GC triplet.

2.2 C-NUCLEOSIDES

High-affinity, sequence-specific recognition of double-stranded DNA by TFOs is important in the selective control of gene expression. Therefore, the replacement of natural nucleosides with non-naturally occurring ones is a frequently applied approach. This enhances the properties, functions and uses of TFOs.

C-glycosidically linked heteroaryl groups that are present in non-naturally occurring nucleosides have incorporated the essential properties of hydrogen bonding and stacking, that are required for DNA interaction. Extensive research has been carried out on the synthesis of C-nucleosides.^{158,201,202,211,213-222}

The analogues should have β -configuration to mimic natural nucleosides. Therefore, a crucial consideration in the synthesis of these C-nucleosides is the stereoselective

incorporation of the heteroaryl base.^{158,211,220,221,223,224} For synthetic procedures that give anomeric mixtures of the *C*-nucleoside,²⁰² the separation of the anomers can be a very challenging step.

2.2.1 Introduction to *C*-nucleosides

C-nucleosides, are a group of nucleosides in which the sugar moiety is linked to a base through a carbon-carbon bond.²²⁵ Some *C*-nucleosides are naturally occurring compounds, e.g., pseudouridine (Ψ), which is isolated from yeast *t*-RNA (certain *t*-RNAs deficient in pseudouridine are incapable of participating in protein synthesis)²²⁵; and showdomycin, an antibiotic with cytotoxic properties (*fig. 2.3*).

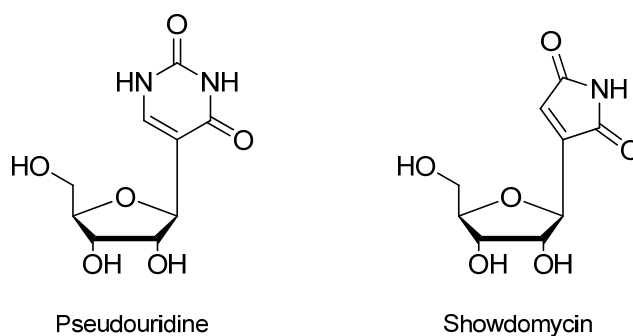


Fig. 2.3: Structures of the naturally occurring *C*-nucleosides Pseudouridine and Showdomycin.

C-nucleosides are much more chemically stable than *N*-nucleosides (glycosidic linkage through a carbon-nitrogen bond), as they possess a hydrolytically and enzymatically more robust sugar-base bond.²²⁶ Replacement of the N atom by C often dramatically changes the properties of the heterocyclic moiety, i.e. the tautomeric populations and acid/base properties of the heterocycles and functional groups (OH, NH₂).²²⁷ *C*-nucleosides have been shown to be suitable candidates for use as building blocks of oligonucleotides for use in gene therapy.²²¹

2.2.2 Synthesis of *C*-Nucleosides²²⁷

Many strategies have been developed for the synthesis of *C*-nucleosides. They can be placed into 5 categories: (1) construction of an aglycon unit on a pre-synthesised carbohydrate moiety; (2) construction of a carbohydrate moiety on a pre-synthesised

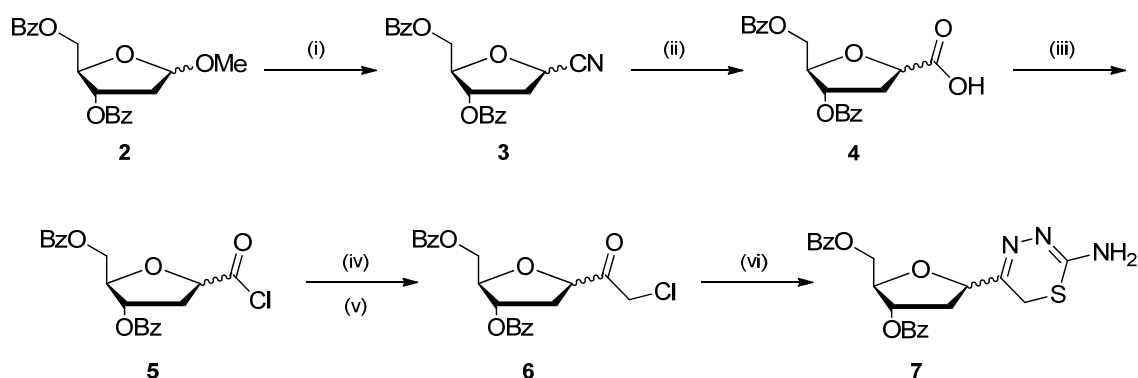
aglycon unit; (3) direct coupling of a carbohydrate moiety with a preformed aglycon unit; (4) modification of existing C-nucleosides; and (5) modifications of variable modules within the nucleosides.

2.2.2.1 Construction of an Aglycon Unit on a Pre-Synthesised Carbohydrate Moiety

2.2.2.1.1 Introduction of the Nitrile Group and subsequent reactions for heterocyclic construction

This approach was the first to introduce a functional group at the anomeric position of the carbohydrate, creating a C-C linkage, followed by a multistep assembly of the heterocycles onto the C-glycosyl derivative.²²⁵

The synthesis reported by Adamo *et al.*²¹⁴ introduces a nitrile group at the anomeric position of a fully protected carbohydrate sugar, 3',5'-dibenzoyl-1'-methoxy-2'-deoxyribose **2**, using trimethylsilyl cyanide, producing a C-C linkage (*scheme. 2.1*). A heterocyclic base was then constructed through numerous reactions starting at the nitrile group in compound **3** to finally give, following cyclisation of **6** with semithiocarbazide, the thiazidine nucleoside **7**, as an equimolar mixture of anomers in good yield.

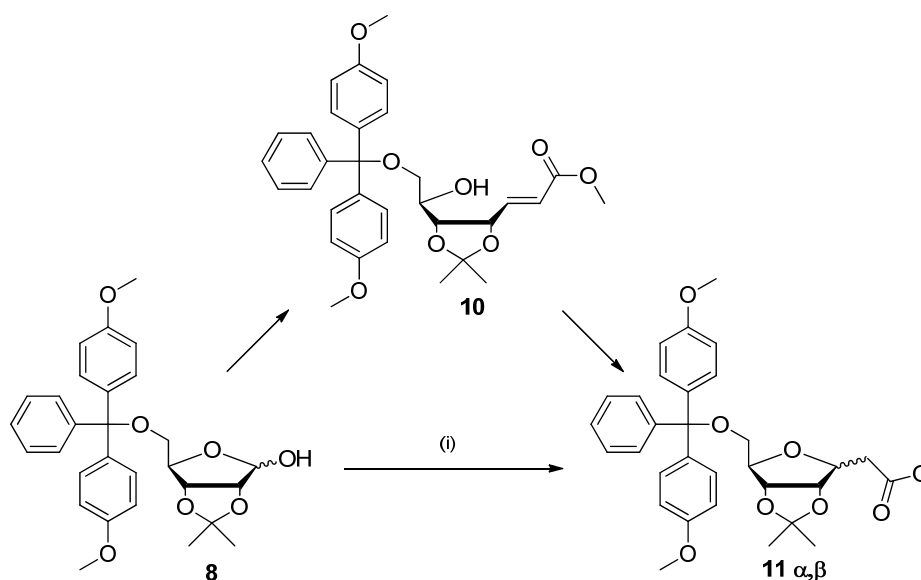


Scheme. 2.1: Adamo *et al.*²¹⁴ synthesis to thiazidine nucleoside **7**. *Reagents and conditions:* (i) DCM, 1.3 eq TMS-CN, 3.3 eq BF₃Et₂O, 0 °C, 1 ½ hrs, 71 %; (ii) dioxane/HCl, 70 °C, 6 hrs, 72 %; (iii) DCM, 5 eq α,α-dichloromethyl methyl ether, reflux, 4 hrs, 99 %; (iv) Et₂O, 3 eq CH₂N₂; (v) HCl gas, 30 mins, 61 %; (vi) methanol, 1.0 eq thiosemicarbazide hydrochloride, reflux, 1 hr, 81 %.

2.2.2.1.2 Wittig-type reactions

Another approach to aglycon construction utilises Wittig-type chemistry, where reaction occurs between a phosphorus ylide and an aldehyde function on the carbohydrate moiety.

The first Wittig-type strategy to produce a C-C linkage was demonstrated by Ohruí *et al.*²¹⁵, their development of this method involved the reaction of 2',3'-*O*-isopropylidene-5'-*O*-trityl-D-ribofuranose, **8** with carboxymethoxymethylenetriphenylphosphorane, **9** (scheme. 2.2). The reaction initially gave **10** which following a Michael-type ring closure gave an anomeric mixture of the C-glycosides **11** α,β (with the more polar isomer dominating in a 3:1 ratio) in quantitative yield. The C-glycosides can then be used to synthesise various nucleoside analogues.



Scheme. 2.2: Ohruí *et al.*²¹⁵ initial synthetic route towards β -D-ribofuranosyl C-glycosides. *Reagents and conditions:* (i) CH₃CN, carboxymethoxymethylenetriphenylphosphorane **9**, reflux overnight.

2.2.2.1.3 Cycloadditions

Cycloaddition reactions utilising Diels-Alder²¹⁶, 1,3-dipolar²¹⁷ and radical species have also been carried out to construct aglycons at the anomeric position of the sugar.²²⁷

2.2.2.2 Construction of a Carbohydrate Moiety upon an Aglycon Unit

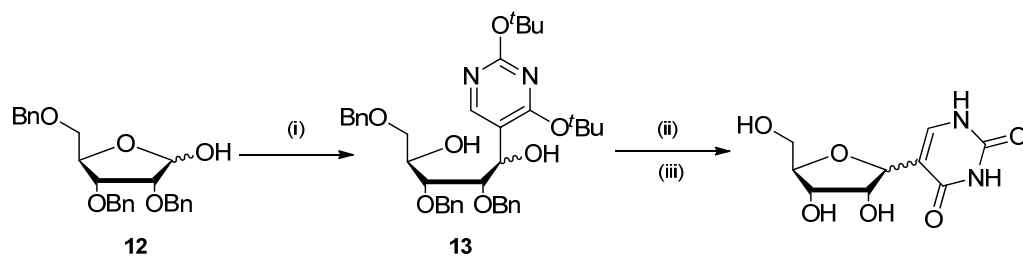
This strategy requires the construction of up to four stereogenic centres, hence it is not commonly used.²²⁷ Very complex synthetic procedures are required in order to obtain the correct stereochemistry about each individual stereogenic centre in the carbohydrate moiety of the nucleosides.

2.2.2.3 Direct Coupling of a Carbohydrate Moiety with a Preformed Aglycon Unit

The most common strategy for the construction of *C*-nucleosides is the direct coupling of a protected sugar moiety with a preformed aglycon nucleophile. There are six different approaches, each involving nucleophilic attack at an electrophilic centre: (1) nucleophilic addition to ribofuranose derivatives; (2) nucleophilic addition to 1',2'-anhydrofuranoses; (3) coupling of nucleophiles with halogenoses; (4) nucleophilic addition to furanolactones; (5) Heck-coupling; and (6) Lewis acid-mediated electrophilic substitution.²²⁷

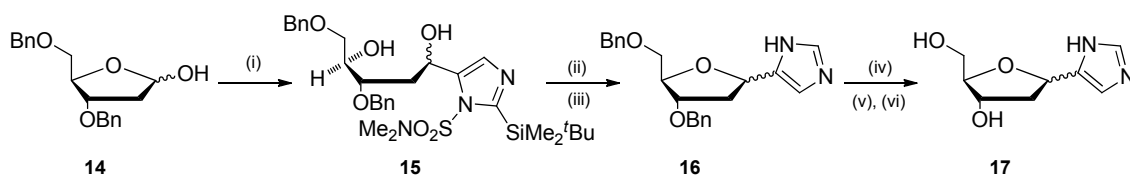
2.2.2.3.1 Nucleophilic addition of organometallic reagents to ribofuranose derivatives

The addition of organometallic reagents with ribofuranose derivatives leads to a mixture of diastereoisomeric diols. Subsequent cyclisation can then proceed in either a non-diastereoselective fashion or diastereoselectively. Cyclisation can take place under acidic conditions,^{219,228,229} as was demonstrated in the synthesis of Pseudouridine reported by Brown *et al.*²¹⁹ (*scheme. 2.3*).



Scheme. 2.3: Synthesis of Pseudouridine and its α -anomer starting from 2',3',5'-tri-*O*-benzyl-D-ribose **12**. *Reagents and conditions:* (i) 2,4-di-(*tert*-butoxy)-5-lithiopyrimidine, -78 °C, THF, 64 %; (ii) HCl, MeOH, r.t, 24 hrs, 90 %; (iii) BCl₃, -78 °C, DCM, 42 %.²¹⁹

The diastereoisomeric diols can also undergo cyclisation under Mitsunobu conditions. This was demonstrated by both Reese *et al.*²⁰² in the synthesis of Pseudouridine and 2-aminopyridine *C*-nucleosides, and by Harusawa *et al.*²²⁰ in the synthesis of imidazole *C*-nucleosides (*scheme. 2.4*).



Scheme. 2.4: Harusawa *et al.*²²⁰ Synthesis of 4(5)-(2'-deoxyribofuranosyl)imidazole starting from 2'-deoxy-3',5'-di-*O*-benzyl-D-ribose **14**. *Reagents and conditions:* (i) 2.5 eq 2-(*tert*-butyldimethylsilyl)-*N,N*-dimethylimidazole-1-sulfonamide, 2.5 eq ⁿBuLi, -50 °C to r.t, THF, 99 %; (ii) 1.5 M HCl, reflux, 1 hr, 92 %; (iii) 1.3 eq Bu₃P, 1.3 eq TMAD, Benzene, r.t, 14 hrs; (iv) Ethyl chloroformate, pyridine, benzene, r.t, 2 hrs, 72 % over steps (iii) and (iv); (v) Pd(OH)₂-C, cyclohexene, quantitative yield; (vi) Hydrazine monohydrate, 90 %.²²⁰

The synthetic procedure developed by Harusawa *et al.*²²⁰ was β -stereocontrolled; the Mitsunobu reaction produced almost exclusively the desired β -anomer (α : β 1:5) from both the *R*- and *S*- epimers. This is thought to be accredited to the extra stabilisation from intramolecular hydrogen bonding between the nitrogen in the imidazole and hydroxyl groups in the sugar, hence enabling control in the glycosylation step (*fig. 2.4*).

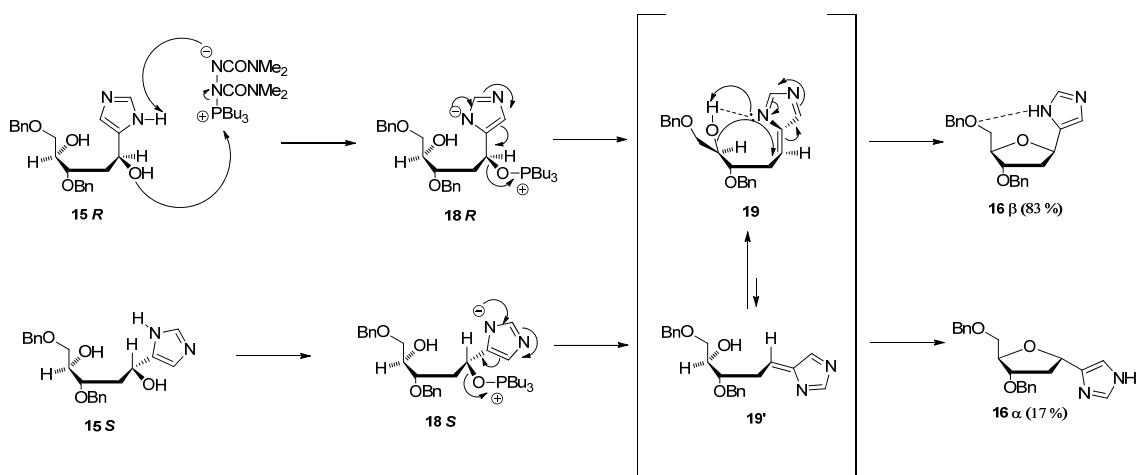
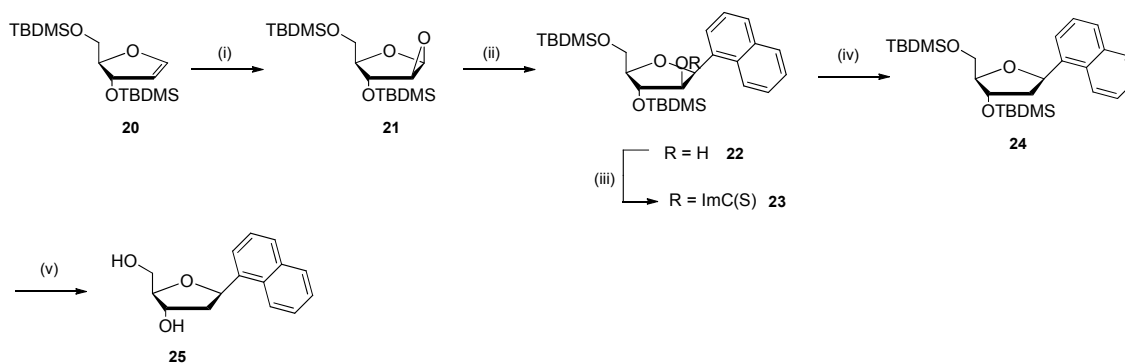


Fig. 2.4: Mechanistic considerations on β -stereoselective glycosylation on the Mitsunobu cyclisation in the synthetic procedure developed by Harusawa *et al.*²²⁰

The procedure for the 2-aminopyridine *C*-nucleoside developed by Reese *et al.*²⁰² however, produced an anomeric mixture (α : β 2:3) that required extra steps to enable separation (discussed in more detail in *section 2.3.1*).

2.2.2.3.2 Nucleophilic addition to 1',2'-anhydrofuranoses

Singh *et al.*²¹⁸ reported the cis-selective opening of β -configured epoxides in 1,2-anhydroarabinose by arylaluminium reagents to give polyaromatic *C*-nucleosides. This provided a facile means of exerting strict stereo-control in *C*-glycosylation reactions (*scheme. 2.5*).



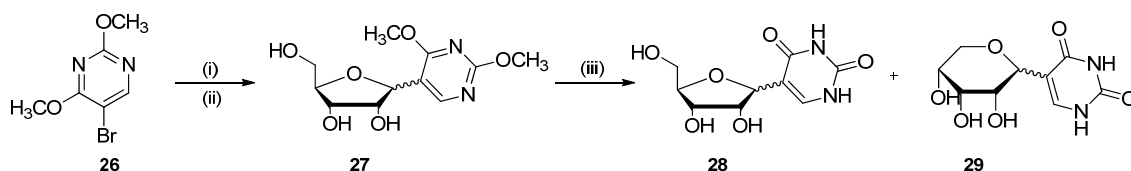
Scheme 2.5: Cis-Selective opening of 1,2-Anhydroarabinose in the synthesis of β -Aryl C-nucleosides by Singh *et al.*²¹⁸ Reagents and conditions: (i) Dimethyldioxirane, CH_2Cl_2 0 °C; (ii) 1-naphthyllithium, AlCl_3 , THF, CH_2Cl_2 , 50 %; (iii) 1,1-Thiocarbonyl diimidazole, CH_2Cl_2 , 90 %; (iv) Tris(trimethylsilyl)silane, Azobisisobutyronitrile, Toluene, 91 %; (v) Tetra-n-butylammonium fluoride, THF, 98 %.

Starting with thymidine, the 3',5'-disiloxane-protected glycal **20** was prepared in two steps,²³⁰ which then underwent epoxidation selectively from the β -side. Cis opening of the epoxide with trinaphthylaluminium was performed followed by a two-step Barton deoxygenation.²³¹ Removal of the protecting groups at the 3'- and 5'- positions of the sugar ring using TBAF in THF afforded the C-nucleoside **25**.

2.2.2.3.3 Coupling of nucleophiles with halogenoses

The oldest method for the construction of the C-glycoside bond was reported by Hurd *et al.*²³² They used Grignard reagents to couple with suitably protected α -glycosyl chlorides to give anomeric mixtures of the coupling product.

Shapiro *et al.*²¹³ were the first to use this approach in the area of C-nucleosides. They reported the synthesis of Ψ -uridine (scheme. 2.6) by the coupling of 2,4-dimethoxypyridine-5-lithium (lithiated **26** at the 5-position) with 2',3',5'-tri-*O*-benzoyl-D-ribofuranosyl chloride (which was prepared from 1'-*O*-acetyl-2',3',5'-tri-*O*-benzoyl- β -ribofuranosyl). The yields however, were very poor (2 % for the natural pseudouridine; 3.5 % for all four pseudouridine isomers), which was due to incomplete hydrolysis of **27**. There was also uncertainty of configuration about the anomeric carbon, and the formation of the pyranosyl isomer, **29**.



Scheme. 2.6: Synthesis of Pseudouridine by Shapiro *et al.*²¹³ *Reagents and conditions:* (i) n-Butyllithium, THF, -75 °C; (ii) 2',3',5'-tri-*O*-benzoyl-D-ribofuranosyl chloride; (iii) Dichloroacetic acid/ water, 100 °C, 4 hrs.

Hoffer's α -chlorosugar²³³ (1-chloro-2-deoxy-3,5-bis[*O*-(*p*-toluoyl)]- α -D-ribofuranose, **30** (*fig. 2.5*) was one of the most widely used building blocks under this approach.

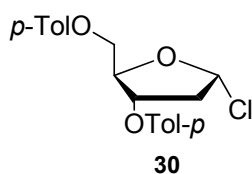


Fig. 2.5: Hoffer's α -chlorosugar.

Typically, reactions would involve the coupling of **30** with organometallic reagents based on various metals (lithium, zinc, mercury, cadmium, mercury). The yields tended to be low in general, and the α -anomer was the dominating product. To obtain the more desirable, naturally configured β -anomer, an additional step was normally required which involved acid-catalysed epimerisation of the α -anomer.^{234,235}

Maebe *et al.*²³⁵ reported the use of trifluoroacetic acid in dichloromethane for the epimerisation of the α -anomer **31**, to give the desired β -anomer **32** (*fig. 2.6*). The combination of protonation of the oxygen within the furan ring, followed by protonation of the furanose ring oxygen by acid allows opening and re-closure to take place. Pure **32** epimerises to **31** under the same conditions, however the rate of epimerisation is significantly lower than that of **31** to **32**.

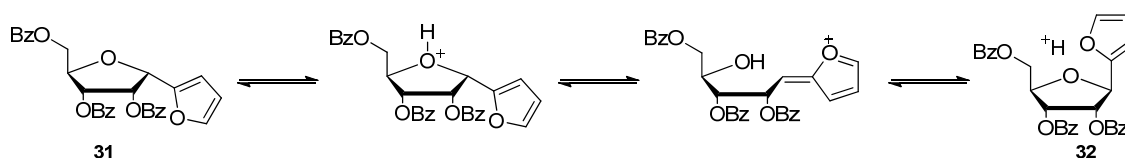


Fig. 2.6: Epimerisation of an α -anomer to give the desired β -anomer, under acidic conditions reported by Maeba *et al*²³⁵.

2.2.2.3.4 Nucleophilic addition to furanolactones

The most frequently used strategy to synthetically construct the C-C glycosidic bond in C-nucleosides is *via* nucleophilic addition of organometallic reagents at low temperatures across a lactone functionality **33** (fig. 2.7). This leads to the formation of a hemiacetal **34** (normally exists as an equilibrium mixture of the open form and two hemiacetal anomers).

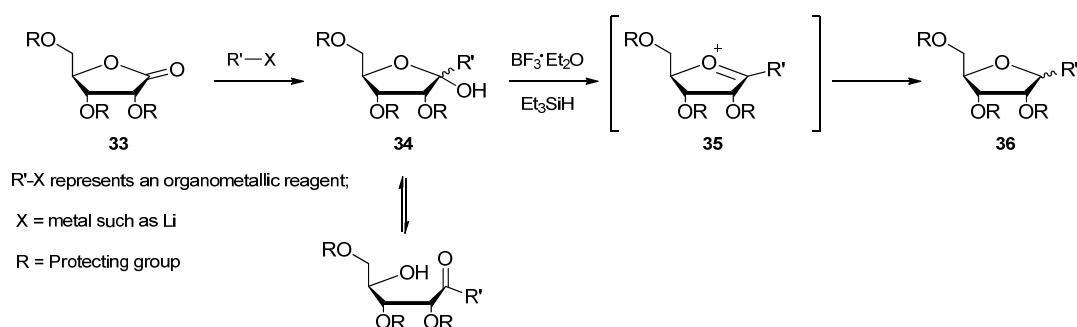
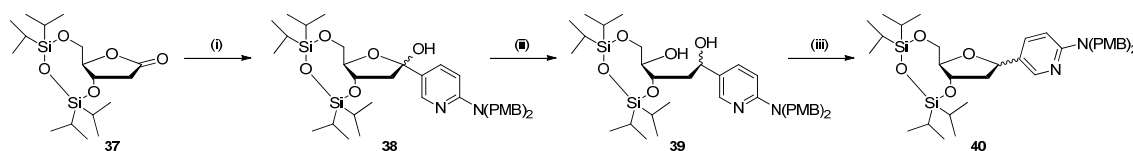


Fig. 2.7: A general representation of nucleophilic addition of an organometallic reagent, R-X across a lactone **33** to give the C-nucleoside, **36**.

The hemiacetal is then reduced to the desired C-nucleoside. A popular reaction for the reduction of the hemiacetal involves initial deoxygenation by a Lewis acid (normally boron trifluoride diethyl etherate), followed by reduction of the oxonium ion intermediate, **35** (usually with triethylsilane) to give rise to the C-nucleoside, **36**. The stereoselectivity normally depends on the nature of the aglycon unit and the protecting groups on the sugar moiety. However, the reduction of the oxonium ion using triethylsilane is known to proceed with good/excellent stereoselectivity giving rise to the β -anomers as either the main or the sole product.^{158,202,211,223,236}

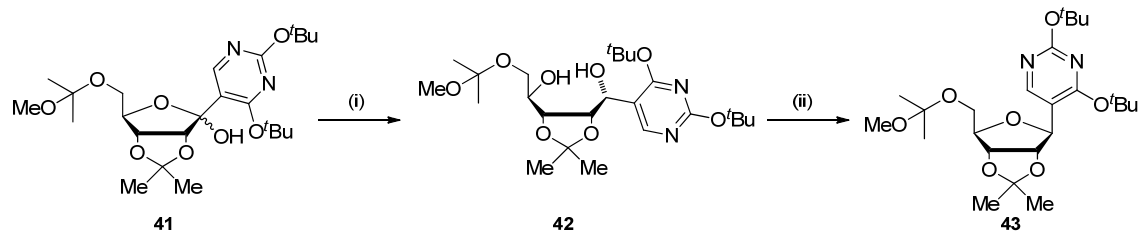
Hildbrand *et al.*^{158,211} and Reese *et al.*²⁰² have utilised this method in their synthetic routes towards the 2-aminopyridine and 3-methyl-2-aminopyridine *C*-nucleosides. Hildbrand *et al.* essentially followed the syntheses of *C*-glycosides by Krohn *et al.*²²³ and Kraus *et al.*²³⁶

Alternatively the hemiacetals, **34** can be reduced to the corresponding diols and then cyclised under Mitsunobu conditions as was demonstrated by Reese *et al.*²⁰² in the synthetic route to 2-aminopyridine. Reduction in their case could be carried out by using either sodium borohydride, or lithium tri-*sec*-butyl borohydride (L-selectride) (*scheme. 2.7*).

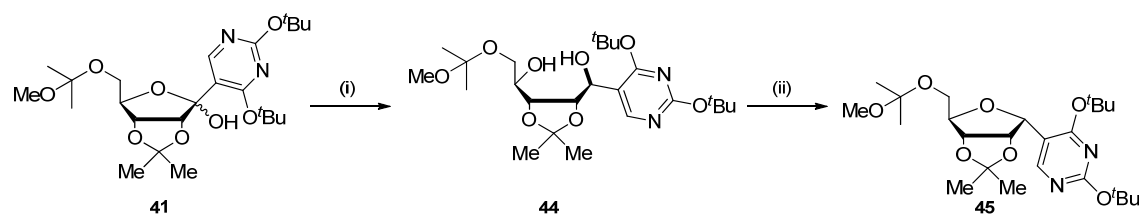


Scheme. 2.7: Synthetic route towards 2-aminopyridine, using nucleophilic addition to a lactone moiety, used by Reese *et al.*²⁰² *Reagents and conditions:* (i) 2-[N,N-bis(4-methoxybenzyl)]aminopyridine-5-lithium, THF, -78 °C, 4 hrs; (ii) NaBH₄, MeOH, 0 °C, 2 hrs or L-Selectride, THF, -78 °C to r.t; (iii) DEAD or DIAD, Ph₃P, THF, 0 °C, 3 hrs.

The hemiacetal can be reduced stereoselectively as was reported by Hanessian *et al.*²³⁷ Reported within this literature was the stereo-controlled synthesis of both α - and β -pseudouridine. Reduction of the hemiacetal, **41** with L-selectride in the presence of ZnCl₂ gave the *altro* isomer, **42** which was then cyclised under Mitsunobu conditions to afford solely the β -anomer, **43** in 70 % yield (*scheme. 2.8*). Whereas reduction of the hemiacetal, **41** with L-selectride in the absence of ZnCl₂ gave the *allo* isomer, **44** which was then cyclised under Mitsunobu conditions to give solely the α -anomer, **45** in 90 % yield (*scheme. 2.9*).



Scheme. 2.8: Stereo-controlled synthesis of β -pseudouridine by Hanessian *et al.*²³⁷ *Reagents and conditions:* (i) L-selectride, ZnCl_2 , $\text{CH}_2\text{Cl}_2/\text{THF}$, $-78\text{ }^\circ\text{C}$ to r.t., 85 %; (ii) DIAD, Ph_3P , THF, 70 %.

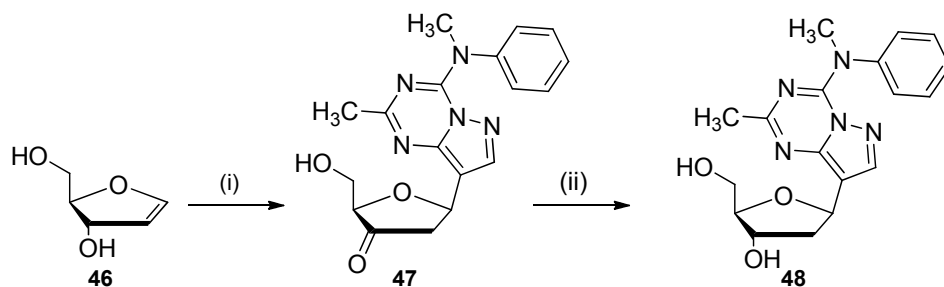


Scheme. 2.9: Stereo-controlled synthesis of α -pseudouridine by Hanessian *et al.*²³⁷ *Reagents and conditions:* (i) L-selectride, $\text{CH}_2\text{Cl}_2/\text{THF}$, $-78\text{ }^\circ\text{C}$ to r.t., 88 %; (ii) DIAD, Ph_3P , THF, 90 %.

2.2.2.3.5 Heck Coupling

The Heck reaction is highly regio- and stereoselective. When applied to glycols the general reaction proceeds *via syn* addition of the organopalladium species to the double bond followed by *syn* elimination of the hydridopalladium salt. The addition reaction is regiospecific owing to the strong polarization of the enol ether double bond; hence the new C-glycosidic bond is formed selectively at the carbon adjacent to the oxygen (the anomeric carbon). In addition, the initial attack occurs from the less sterically hindered face of the glycol ring. Therefore, the stereoselectivity can be controlled by the use of bulky or small groups at the 3' position of the glycol, along with the use of bulky or small ligands in the palladium mediated cross coupling reaction.

The Heck reaction has been important in the construction of the C-C glycosidic bond in the synthesis of 2'-deoxy-C-nucleosides; the first reported use of this method was by Arai *et al.*²³⁸ More recently, Raboisson *et al.*²²⁴ reported the stereocontrolled synthesis of a C-nucleoside to give solely the β -anomer (*scheme. 2.10*).



Scheme. 2.10: Stereocontrolled synthesis of the C-nucleoside, **48** starting with the glycal, **46** in a palladium mediated cross coupling reaction by Raboisson *et al.*²²⁴ *Reagents and conditions:* (i) 8-Iodo-2-methyl-4-(*N*-methyl-*N*-phenylamino)pyrazolo-[1,5-*a*]-1,3,5-triazine, bis(dibenzylideneacetone)Pd(0), Ph₃As, TEA, 75%; (ii) sodium triacetoxyborohydride, 92%.

A completely unprotected glycal, **46** was used. However, the use of the bulky palladium ligands (triphenylarsine) made the organopalladium reagent (obtained via transmetallation) a much hindered complex. Hence, it was probably not able to attack the olefinic glycal through the face bearing the 3'-hydroxyl group. Therefore, *syn* addition occurred on the sterically most open face of the glycal (α face) to give the adduct, **49** (*fig. 2.8*).

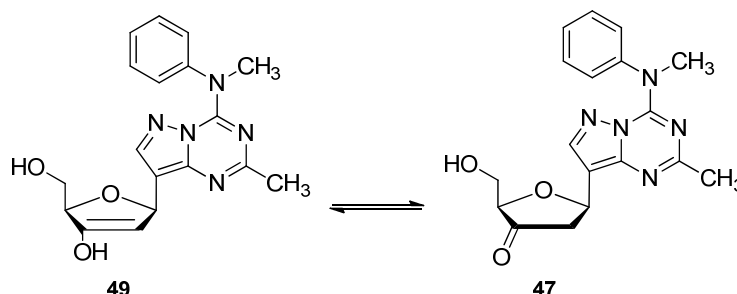


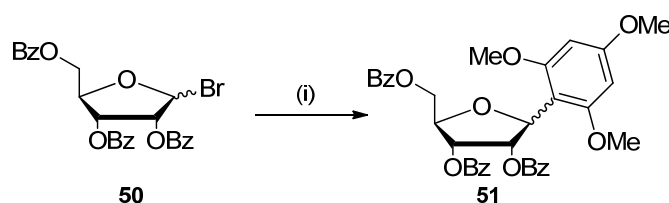
Fig. 2.8: The keto-enol tautomerism equilibrium that exists between compound **49** and compound **47**.

This adduct is in tautomeric equilibrium with compound **47**, which then undergoes stereospecific reduction on the 3'-keto group with sodium triacetoxyborohydride to yield the 2'-deoxyribofuranosyl C-nucleoside, **48**. When K-selectride is used in place of the sodium triacetoxyborohydride reagent, reduction is permitted by the less sterically hindered face of the sugar ring, forming the stereoisomer of **48**, where the 3'-hydroxyl group resides on the *endo*-face of the furanose.

2.2.2.3.6 Lewis-acid mediated electrophilic substitutions

A simple route to *C*-nucleosides is *via* the coupling of a preformed carbohydrate moiety with a (hetero)aryl, the reaction is catalysed by a Lewis acid (Friedel-Crafts type electrophilic substitution of the aromatic ring by the glycon). Although a simple process, it is associated with disadvantages of poor regioselectivity of the aglycon attack, and only modest stereoselectivity of the anomeric C-C bond formation. The nature of the aglycon controls the regioselectivity of the reaction. The interaction of the protected carbohydrate, the aglycon unit and the Lewis acid affect the diastereoselectivity and yield. This Friedel-Crafts method is applicable to only some electron-rich arenes and heterocycles, in some cases it is accompanied by double arylation of the sugar to form the undesired 1,1-diaryl alcohols.

Kalvoda *et al.*²²² used this approach in the initial steps for the synthesis of Showdomycin (*scheme. 2.11*). Condensation of benzylated bromose, **50** with 1,3,5-trimethoxybenzene in the presence of the Lewis-acid, zinc oxide gave the *C*-nucleoside, **51** with α : β C-C bond formation ratio of 1:1.



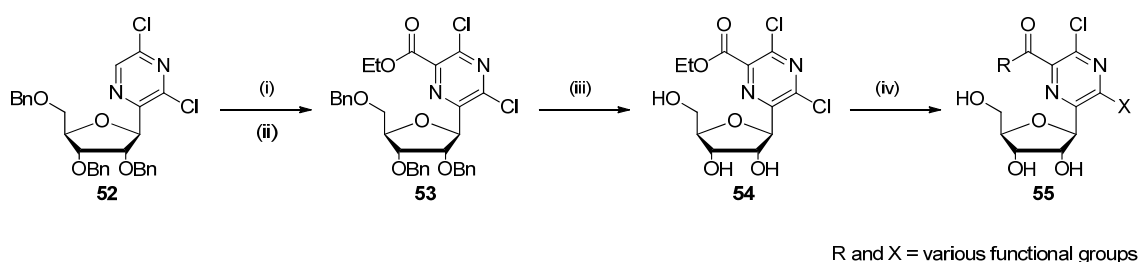
Scheme. 2.11: Initial steps for the synthesis of Showdomycin by Kalvoda *et al.*²²² using a Friedel-Crafts-type approach. *Reagents and conditions:* (i) 1,3,5-trimethoxybenzene, ZnO, benzene, r.t.

2.2.2.4 Modification of existing *C*-nucleosides

This approach involves the modifications of the functional groups that exist within natural or synthetic *C*-nucleosides. The modifications are either on the aromatic or the sugar moiety, or both; the C-C glycosidic bond is kept intact during all transformations.

This approach involves the stereoselective synthesis of a common intermediate that can, through further reactions give a vast range of *C*-nucleosides with different functions and

properties. Walker *et al.*²³⁹ reported the synthesis of a series of pyrazonic acid C-nucleosides (*scheme. 2.12*).



Scheme. 2.12: Synthetic route developed by Walker *et al.*²³⁹ to produce a variety of functionalised pyrazine C-nucleosides **55**. *Reagents and conditions:* (i) LTMP, THF, -94 °C; (ii) Ethyl cyanoformate; (iii) BCl₃, DCM, -70 °C; (iv) Various conditions.²³⁹

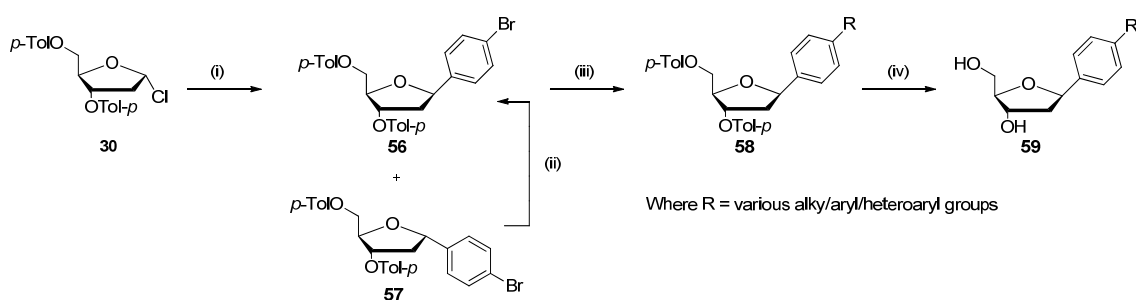
The fully protected pyrazine C-nucleoside, **52** was selectively lithiated by lithium 2,2,6,6-tetramethylpiperidide (LTMP) at low temperature. Electrophilic esterification of the lithiated derivative using ethyl cyanoformate, and subsequent debenzylation gave the common intermediate, ethyl 3,5-dichloro-6-(β-D-ribofuranosyl)pyrazine-2-carboxylate, **54**. This was then subjected to a variety of reaction conditions to give a series of functionalised pyrazine C-nucleosides **55** which were tested for antiviral activity.

There exist many synthetic procedures in which the stereoselectivity around the glycosidic bond is only moderate; hence controlled epimerisation is highly desirable. Jiang *et al.*^{240,241} have reported efficient catalysed epimerisation ($\alpha \rightarrow \beta$ and $\beta \rightarrow \alpha$) of C-nucleosides containing electron donating groups using trifluoroacetic acid in dichloromethane. C-nucleosides containing electron withdrawing groups can undergo epimerisation in the presence of combination catalysts TFA-benzenesulfonic acid (5:1 molar ratio), in dichloromethane. Trifluoroacetic acid is non-oxidative; benzenesulfonic acid on its own is oxidative and causes substrate decomposition and dehydration. For both procedures, no dehydration, carbonisation or decomposition was observed. Hence, for synthetic routes that have shown only moderate stereoselectivity for the desired anomer, these procedures are useful tools to improve the stereoselectivity and hence the overall yield of the C-nucleoside.

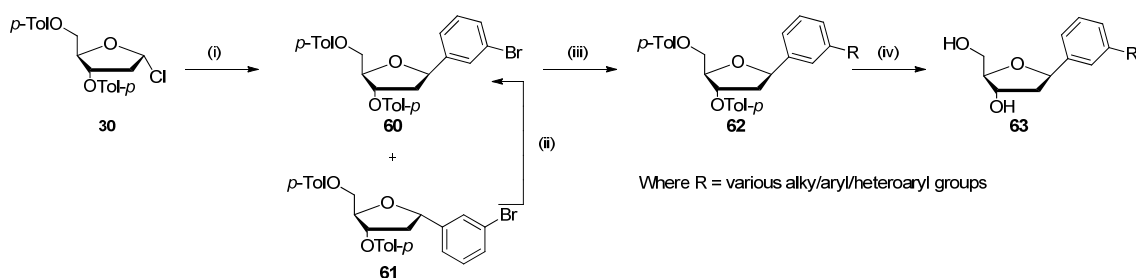
2.2.2.5 Modular approaches to C-nucleoside synthesis

This method can be used in the construction of libraries of C-nucleosides. Once the synthetic route to the underlying framework of the nucleoside has been developed, the modules that can be varied are: the sugar moiety, the aromatic aglycon and the functional groups that exist.

Hocek *et al.*²⁴²⁻²⁴⁶ have developed many modular approaches to C-nucleosides. They reported the synthesis of 4- and 3- substituted phenyl C-nucleosides based on a facile synthesis of protected 4- or 3- bromophenyl C-nucleoside intermediates respectively, and their standard palladium mediated cross-coupling reactions with diverse organometallics (*scheme. 2.13* and *2.14*).²⁴²



Scheme. 2.13: Hocek *et al.*²⁴² modular approach towards C-nucleosides. *Reagents and conditions:* (i) 1,4-dibromobenzene, Mg, ethylene bromide, THF, 65 °C then r.t; (ii) Benzenesulfonic acid-TFA, DCM, 40 °C, 24 hrs; (iii) RM (Organometallic reagent), Pd(PPh₃)₄; (iv) NaOMe, MeOH, r.t, overnight.



Scheme. 2.14: Hocek *et al.*²⁴² modular approach towards C-nucleosides. *Reagents and conditions:* (i) 1,3-dibromobenzene, Mg, ethylene bromide, THF, 65 °C then r.t; (ii) Benzenesulfonic acid-TFA, DCM, 40 °C, 24 hrs; (iii) RM (Organometallic reagent), Pd(PPh₃)₄; (iv) NaOMe, MeOH, r.t, overnight.

1,4- and 1,3-dibromobenzene were converted to mono-Grignard reagents and coupled separately with a fully protected halogenose, **30** (Hoffer's α -chlorosugar).²³³ The C-nucleoside **56/60** was then subjected to a palladium mediated cross-coupling reaction with various organometallic reagents (either Suzuki-Miyaura conditions, Stille conditions for **56** only, or the use of organozinc or organoaluminium reagents for **56** only), hence generating a large series of diverse novel C-nucleosides.

In conclusion, there exists a variety of methods for the construction of C-nucleosides. These are constantly being improved and altered; there is no generally applicable and efficient method. The method of choice normally depends on the nature of the aglycon and the glycal. The majority of the earlier methods for synthesising C-nucleosides involve many steps and hence resulted in rather low yields and poor stereoselectivity.²²¹ The approach of direct attachment of a pre-formed aglycon unit to an appropriate carbohydrate overcomes the drawbacks of these initial approaches and hence was the strategy of choice in developing the 3-methyl-2-aminopyridine C-nucleoside. However, as mentioned above, only a limited range of electron-rich aromatic moieties can be used.

2.3 SYNTHESIS OF 3-METHYL-2-AMINOPYRIDINE NUCLEOSIDE (^{Me}P)

2.3.1 Background work

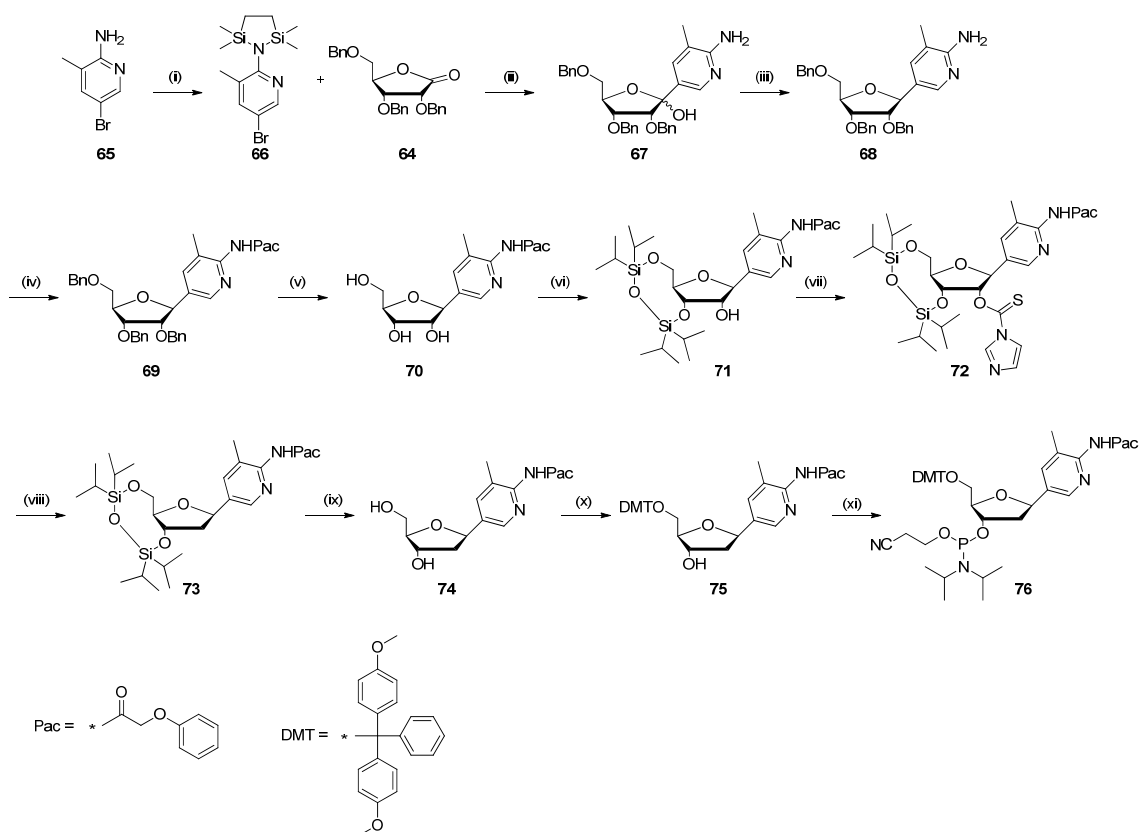
To date there have been three previously published reports relating to the synthesis of ^{Me}P (**1**). Hildbrand *et al.*¹⁵⁸ (*scheme. 2.15*) started with the readily available starting materials, ribonolactone, **64** (obtained in 3 steps from D-ribose) and the protected 2-amino-5-bromo-3-methylpyridine, **66** (obtained in 1 step from **65** in *scheme 2.15*). The coupling step followed the syntheses of C-glycosides by Kraus *et al.*²³⁶ and Krohn *et al.*²²³; bromide-lithium exchange in **66** with *n*-BuLi at -75 °C and *in situ* reaction with lactone **64** afforded a mixture of hemiacetals, **67**. The hemiacetals were subsequently reduced with excess triethylsilane, borontrifluoride diethyletherate to give exclusively the β -configured nucleoside, **68**; the exocyclic amino protection had been cleaved during the coupling reaction. Next, protection of the exocyclic amino function with

phenoxyacetic anhydride, followed by debenzylation with Lewis acid boron tribromide provided the *N*-protected ribo-*C*-nucleoside, **70**. The 3' and 5' hydroxyl groups were then selectively protected with 1,3-dichloro-1,1,3,3-tetraisopropylidisiloxane (Markiewicz reagent) to afford **71**. A two-step Barton²³¹ reaction followed; activation of the 2'-hydroxyl group with thiocarbonyl diimidazole followed by homolytic reductive cleavage of the C-O bond using azoisobutyronitrile (AIBN) and tris(trimethylsilyl)silane gave the 2'-deoxy fully protected nucleoside **73**. Desilylation of **73** with tetra-*n*-butylammonium fluoride (TBAF) afforded the *N*-protected *C*-nucleoside, **74**. This was then protected at the 5' position of the sugar ring using 4,4'-dimethoxytrityl chloride in pyridine, to give **75**. This was then subsequently phosphitylated at the 3' position using *N,N*-diisopropylchlorophosphoramidite and *N,N*-diisopropylethylamine to afford the final *N*-protected monomer, **76**. The free nucleoside ^{Me}P (**1**) was obtained by treatment of **74** with 25 % aqueous ammonia, in 63 % yield.

The monomer **76** was obtained in 14 steps in *ca.* 9.7 % overall yield (starting from *step (ii) in scheme. 2.15*). The free nucleoside ^{Me}P (**1**) was obtained in 13 steps in *ca.* 10.1 % overall yield (starting from *step (ii) in scheme. 2.15*). The yield to obtain the protected base, **66** was 69 %. Importantly, the glycosylation step (*step (iii)*) was β -stereoselective.

One problem encountered by Hildbrand *et al.*¹⁵⁸, was the partial acetylation of the Pac protected amino groups in the residues of ^{Me}P during the capping step of automated oligonucleotide synthesis. These acetyl groups could be removed by methylamine treatment (40 % in water, 70 °C, 24 hrs). This initial problem was then overcome by the use of phenoxyacetic anhydride as the capping agent instead of acetic anhydride, which allowed for efficient post synthetic deprotection under standard ammonolysis conditions (25 % NH₃, 55 °C, ~ 18 hrs).

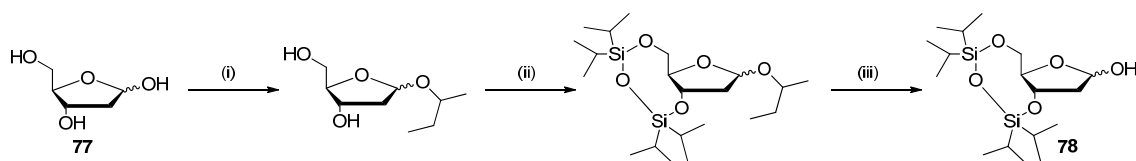
This *C*-nucleoside derivative was incorporated as protonated Cytidine equivalents in homopyrimidine oligonucleotides. T_m measurements indicated that over the pH range 6.0-8.0, the oligonucleotides containing ^{Me}P had a higher affinity to double-stranded DNA than those containing 5-methylcytidine (^{Me}C). This increased stability was most pronounced above pH 7.0. ^{Me}P containing oligos displayed much lower pH dependence and sequence composition effects than those containing ^{Me}C, in the pH range 6.0-8.0; this was a direct consequence of the higher pK_a.



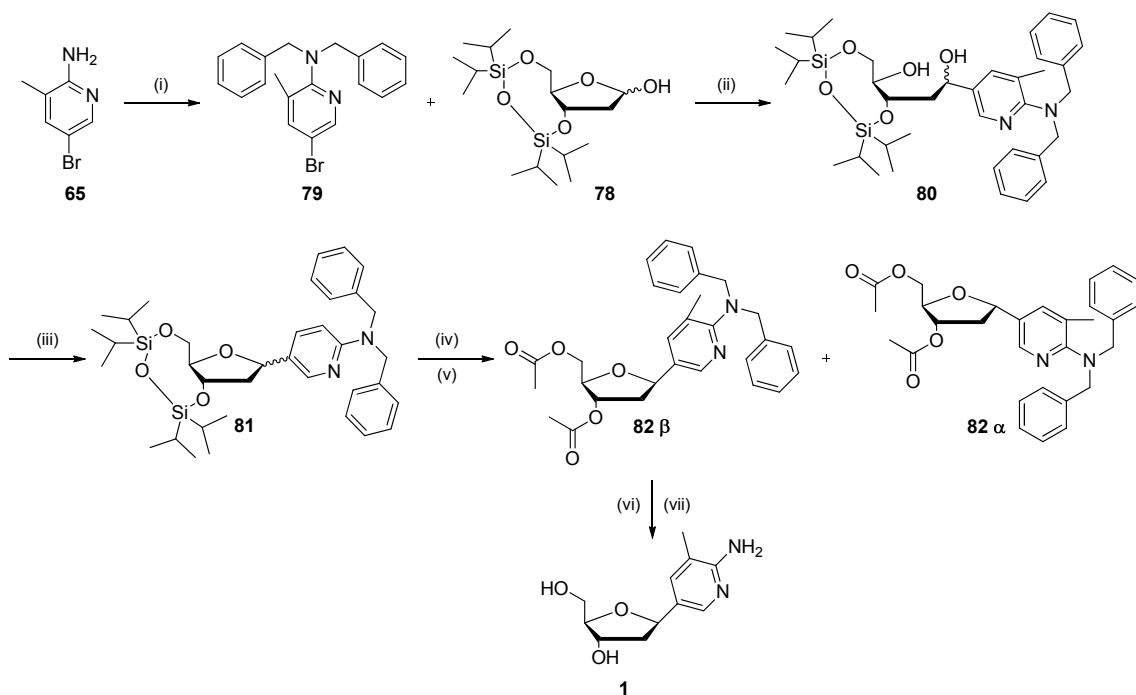
Scheme. 2.15: Hildbrand *et al.*¹⁵⁸ synthesis to *N*-protected ^{Me}P monomer, **76**. *Reagents and conditions:* (i) 1 eq 1,2-bis[(dimethylamino)dimethylsilyl]ethane, 0.7 mol % ZnI₂, 140 °C, 15 hrs, 69 %; (ii) 1.5 eq **66**, 1.5 eq *n*-BuLi, THF, -75 °C, 3 hrs, then 1 eq **64**, -75 °C → 0 °C, 4 hrs; (iii) 5 eq Et₃SiH, 5 eq BF₃·Et₂O, DCM, -75 °C → r.t., overnight, 77 % over steps (ii) and (iii); (iv) 3 eq phenoxyacetic anhydride, pyridine, r.t, 4 ½ hrs, 67 %; (v) 0.4 M BBr₃ in DCM, -75 °C, 5 hrs, 81 %; (vi) 1.2 eq (Cl(*i*Pr)₂Si)₂O, pyridine, r.t, 4 hrs, 84 %; (vii) 2.5 eq (Im)₂CS, DMF, 40 °C, 4 hrs, 74 %; (viii) 0.1 eq AIBN, 1.15 eq (Me₃Si)₃SiH, toluene, 80 °C, 5 hrs, 67 %; (ix) 2 eq Bu₄NF, THF, r.t, 1 ½ hrs, 93 %; (x) 1.2 eq DMTCl, pyridine, r.t, 3 ½-8 hrs, 74 %; (xi) 1.5 eq *N,N*-diisopropyl chlorophosphoramidite, 3 eq *N,N*-diisopropylethylamine, THF, r.t, 2 hrs, 81 %.

Reese *et al.*²⁰² reported an alternative synthetic route to ^{Me}P (**1**), which was based on the commercially available starting material 2'-deoxy-D-ribose, **77**. This route has not been optimised. The fully protected 2'-deoxy lactol **78**, which was obtained in three steps (*scheme. 2.16*), was coupled with the dibenzyl protected lithiated pyridine, **79** at low temperatures to give a mixture of diols, **80** (*scheme. 2.17*). Following Mitsunobu cyclisation, the fully protected *C*-nucleoside, **81** (β:α 3:2) was obtained. The chromatographic separation of the α- and β-anomers did not prove feasible at this stage. Hence, the mixture had to be converted to the corresponding 3',5'-diacetate mixture, **82**, using a two-step process; first treatment with tetraethylammonium fluoride (TEAF) in

acetonitrile solution followed by reaction with acetic anhydride in pyridine. The pure β - and α -anomers were isolated in this way after chromatographic separation, in 47 % and 7 % yields respectively. Treatment of the β -diacetate, **82** with ammonium cerium (IV) nitrate (CAN) in wet acetonitrile, followed by treatment with alcoholic methylamine, gave the free nucleoside **MeP (1)** in 58 % yield over the two steps.



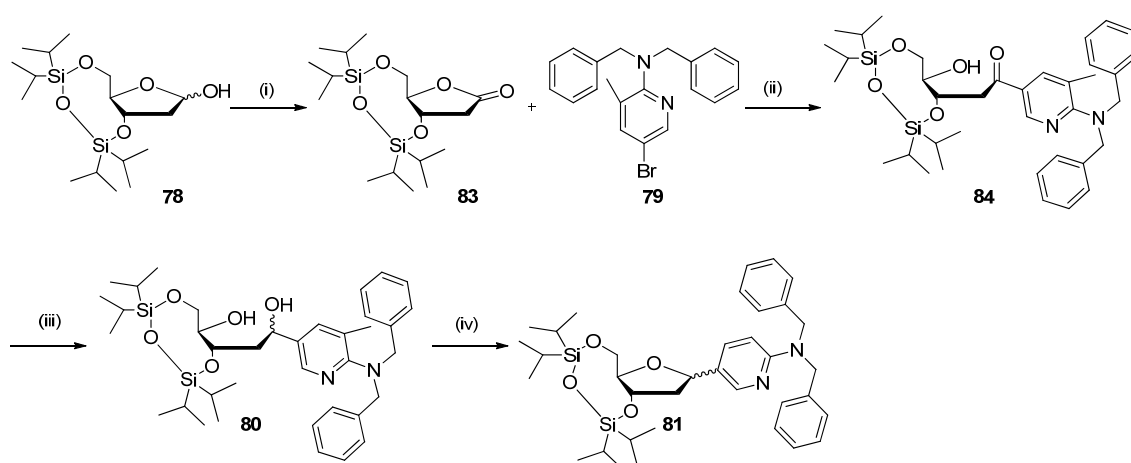
Scheme. 2.16: The synthetic route to the fully protected 2'-deoxy lactol **78** by Reese *et al.*²⁰² *Reagents and conditions:* (i) butan-2-ol, HCl, r.t, 25 mins; (ii) 1,3-dichloro-1,1,3,3-tetraisopropylidisiloxane, imidazole, THF, r.t, 90 mins; (iii) TFA, H₂O, CH₂Cl₂, -15 °C, 20 mins, 90 % over the 3 steps.



Scheme. 2.17: The synthetic route to **MeP (1)** by Reese, *et al.*²⁰² *Reagents and conditions:* (i) Benzyl bromide, NaH, *N,N*-dimethylacetamide (DMA), THF, 91 %; (ii) *n*BuLi, *n*-hexane, THF, -78 °C, 40 mins then **78**, THF, -78 °C, 1 hr then -78 °C→0 °C over 3 hrs; (iii) DIAD, Ph₃P, THF, 0 °C, 3 hrs, 41.5 % over steps (ii)-(iii); (iv) Et₄NF, MeCN, r.t, 90 mins; (v) Ac₂O, 1-methylimidazole, pyridine, r.t, 54 % over steps (iv)-(v); (vi) CAN, MeCN-H₂O (98:2 v/v), r.t, 25 mins; (vii) MeNH₂, EtOH, r.t, 16 hrs, 58 %.

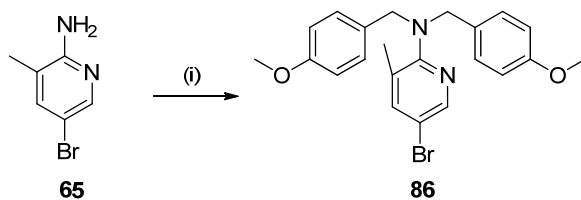
To summarise, the route provided by Reese *et al.* is a 10-step synthesis to the free β C-nucleoside **MeP (1)** in *ca.* 7.8 % overall yield (starting from *step (ii) in scheme. 2.17*).

The overall yield to obtain the fully protected lactol **78** was 90 %, and that to obtain the protected base **79** was 91 %. The Mitsunobu cyclisation reaction in this case was not strongly β -stereoselective. Reese *et al.* suggested that an anomeric mixture that was richer in the β -anomer could possibly be achieved by starting the coupling reaction of the glycal with aglycon, from the protected lactone, **83** (scheme. 2.18). Following this, reduction of the hydroxy-ketone, **84** with L-Selectride to give the diols **80**, and then Mitsunobu cyclisation should produce a higher β : α ratio of compound **81**.



Scheme. 2.18: Alternative route suggested by Reese *et al.*²⁰² to improve the β : α ratio of compound **81**.
Reagents and conditions: (i) pyridinium chlorochromate, DCM, r.t, 16 hrs; (ii) *n*BuLi, *n*-hexane, THF, -78 °C, 40 mins then **83**, THF, -78 °C, 1 hr then -78 °C \rightarrow 0 °C over 3 hrs; (iii) L-Selectride, THF, -78 °C \rightarrow r.t; (iv) DIAD, Ph₃P, THF, 0 °C, 3 hrs.

The third synthesis of ^{Me}**P** monomer was documented in the published thesis by Dr. V. E. C. Powers.²⁰¹ The synthetic scheme developed, essentially followed the work reported by Hildbrand *et al.*^{158,211} via the perbenzylated lactone,^{247,248} **64** and protected 2-amino-5-bromo-3-methylpyridine, **86** (scheme. 2.15). Unlike Hildbrand *et al.* however, the oxidation step of the perbenzylated lactol, **85** (scheme. 2.21) did not follow that by Timpe *et al.*²⁴⁹ This was due to the bad odorous dimethylsulfide being formed as a by-product, and the lactone product being contaminated. Instead, the oxidising agent, tetrapropylammonium perruthenate (TPAP) was used along with the oxidant *N*-methylmorpholine-*N*-oxide to yield the lactone in 85 %. Also, 2-amino-5-bromo-3-methylpyridine was protected with paramethoxy benzyl groups to give **86** in 78 % yield (scheme. 2.19), as did Reese *et al.*²⁰², rather than the disilyl protecting group which was used by Hildbrand *et al.*¹⁵⁸



Scheme. 2.19: Synthetic route to 5-Bromo-2-[*N,N*-bis(4-methoxybenzyl)]amino-3-methylpyridine, **86**.

Reagents and conditions: (i) 2.5 eq 4-methoxybenzyl chloride, 2.5 eq NaH, DMF, r.t, 6 hrs, 78 %.²⁰²

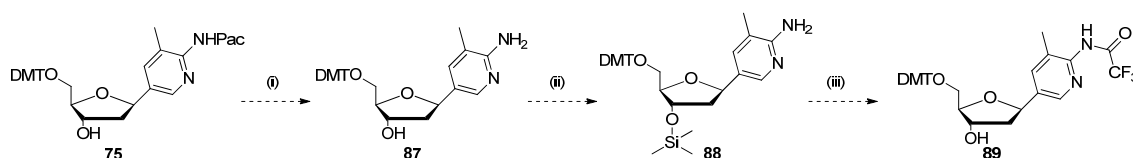
To summarise, the synthetic route developed by Dr. Powers²⁰¹ was an adaptation of the route developed by Hildbrand *et al.*¹⁵⁸; a 15-step scheme to the monomer **76**, with overall yield of *ca.* 5.6 % (starting from the coupling of the lactone, **64** with the fully protected base, **86**). The overall yield to obtain the perbenzylated lactone **64** was *ca.* 81 % and that to obtain the fully protected base was 78 %. The lower overall yield of the synthesis compared to that obtained by Hildbrand *et al.*¹⁵⁸ was attributed to the low yield obtained during the tritylation reaction (39 %), due to the combination of impure starting material, and the 4,4'-dimethoxytrityl chloride reagent that was used had decomposed.

2.3.2 The approaches towards ^{Me}P carried out during this PhD

Discussed herein are the routes that were attempted and accomplished to achieve the desired monomer, ^{Me}P.

The initial aim of this research was to follow the synthetic route to ^{Me}P monomer developed by Dr. Powers²⁰¹, which was based on the routes developed by Reese *et al.*²⁰² and by Hildbrand *et al.*^{158,211} The final monomer would be incorporated with trifluoroacetyl (TFA) protection instead of the phenoxyacetyl, at the exocyclic amino moiety. The reason for this was that, trifluoroacetamide is easily cleaved after oligonucleotide synthesis, under mildly basic conditions (methanolic ammonia, r.t)²⁵⁰ in comparison to phenoxyacetamide which requires harsher conditions (25 % NH₃, 55 °C, ~ 18 hrs). Capping cannot be used during oligonucleotide synthesis, however. This is because the acetic anhydride used would replace the TFA group with an acetyl group which is too stable, and requires harsh conditions to be hydrolysed, as was observed by Hildbrand *et al.*¹⁵⁸ The ^{Me}P monomer would then be incorporated into various TFOs for biophysical analysis to investigate the interactions involved with duplex targets.

Initially, the TFA group was planned to be incorporated towards the end of the synthetic scheme i.e. the phenoxyacetyl protection in **75** from *scheme. 2.15* would be cleaved after *step (x)*, followed by blocking the 3' position of the sugar with trimethylsilyl and then protection of the exocyclic amino function with TFA. The TMS group would be cleaved in transit during aqueous work-up to give **89** (*scheme. 2.20*).

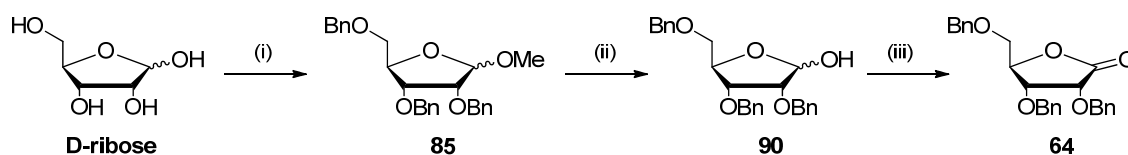


Scheme. 2.20: Potential method to incorporate the required TFA protection into the ^{Me}**P** monomer. *Reagents and conditions:* (i) 25 % NH₃, 55 °C, 18 hrs; (ii) TMSCl, Pyridine, r.t.; (iii) trifluoroacetic anhydride, DCM, 0 °C → r.t.^{158,251}

By doing this the synthetic route would be increased to 18-steps. In an attempt to keep the route as short as possible, it was proposed that the TFA protecting group be added to **68**, post removal of the PMB groups (*see scheme. 2.15*). However, TFA has been known to be hydrolysed under either aqueous or mild acidic/basic conditions²⁵², hence in doing this two potential problems could arise; (i) the TFA group could potentially be cleaved by the highly acidic HBr generated during the debenzylation step using the Lewis acid BBr₃ or; (ii) it could also potentially be cleaved under the highly basic conditions of TBAF during the removal of TIPS. Problems were encountered mid-scheme, and hence the exocyclic amino protecting group had to be altered (discussed in further detail below).

The perbenzylated lactone, **64** was synthesised from D-ribose in three steps (*scheme. 2.21*). The initial part of the first step^{247,253} involved protection of the anomeric position with acidic methanol to form the methyl glycoside, necessary to keep the sugar moiety in the furanose conformation. The second part of this step was complete benzylation at the 2'-, 3'-, and 5'-positions of the anomeric mixture using benzyl bromide and sodium hydride; this one-pot reaction gave the fully protected glycoside, **85** in *ca.* 90 % yield. The methyl group in **85** was then hydrolysed under acidic conditions²⁰¹ to give the perbenzylated lactol **90** in *ca.* 66 % yield. Catalytic oxidation²⁴⁸ of the lactol, using *ca.*

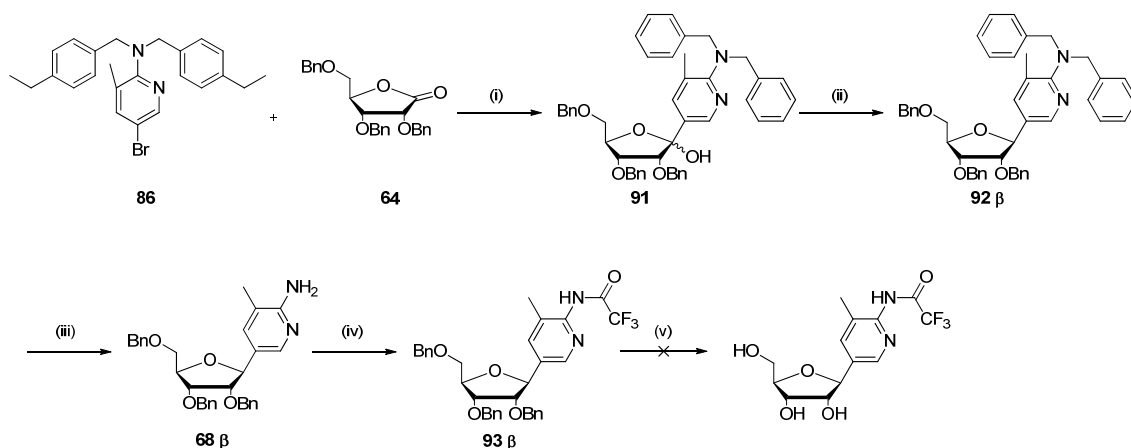
5 mol% of TPAP and 1.6 equivalents N-methylmorpholine-N-oxide with 4 Å powdered molecular sieves gave the fully protected lactone, **64** in *ca.* 83 % yield.



Scheme. 2.21: Synthetic route to 2',3',5'-tri-*O*-benzyl-D-ribo-1,4-lactone, **64**. *Reagents and conditions:* (i) 0.06 eq CH₃COCl, MeOH, r.t, 24 hrs, then 6 eq Benzyl Bromide, 6 eq NaH, DMF, 15 hrs, 90 %; (iii) AcOH (80 % in water), c.H₂SO₄, reflux, 120 °C, 30 mins, 66 %; (iv) 5 mol % TPAP, 1.6 eq NMO, 4 Å powdered molecular sieves, DCM, 0 °C→r.t, 4 hrs, 83 %.^{201,247,248,253}

The PMB protected 2-amino-5-bromo-3-methylpyridine, **86** (*scheme. 2.19*) was afforded in *ca.* 76 % yield, using an excess of paramethoxybenzyl chloride and sodium hydride in DMF at room temperature.²⁰²

Bromide-lithium exchange in **86** with *n*BuLi at -78 °C and *in situ* reaction with lactone **64** furnished a mixture of hemiacetals, **91** which were subsequently reduced with excess triethylsilane and boron trifluoride diethyletherate to give the fully protected β-configured nucleoside, **92 β** (*scheme. 2.22*).¹⁵⁸ Acidic cleavage of the PMB groups with trifluoroacetic acid^{201,202} provided **68 β** in *ca.* 59 % over the three steps. The exocyclic amino function was then protected with trifluoroacetic anhydride yielding **93 β** quantitatively.²⁵⁴ This was then subjected to debenzylation under Lewis acid conditions of boron tribromide at low temperature.¹⁵⁸ Unfortunately, the trifluoroacetyl group cleaved along with the benzyl groups to afford the free ribo-*C*-nucleoside, **94 β** (*fig. 2.9*). The trifluoroacetyl group was thought to have been hydrolysed by the highly acidic HBr formed during this reaction. This route was unsatisfactory due to the observed lability of the trifluoroacetyl protecting group, and therefore discontinued.



Scheme. 2.22: Attempt 1 for the synthesis of ^{Me}P monomer. *Reagents and conditions:* (i) 1.03 eq *n*BuLi, THF, - 78 °C, 1 ½ hrs, then 1.07 eq **64**, THF, - 78 °C→0°C, 1 hr; (ii) 6 eq Et₃SiH, 6 eq BF₃·Et₂O, DCM, - 78 °C→r.t, 16 hrs; (iii) DCM, trifluoroacetic acid (1:1), 20 hrs, 59.1 % over steps (i)-(iii); (iv) 1.1 eq trifluoroacetic anhydride, pyridine, r.t, 17 hrs, quant.; (v) 4.5 eq BBr₃ in DCM, - 78 °C, 4 hrs, then MeOH, 0°C→r.t, overnight.^{158,201,202,254}

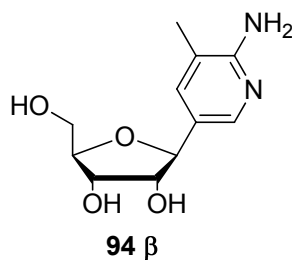
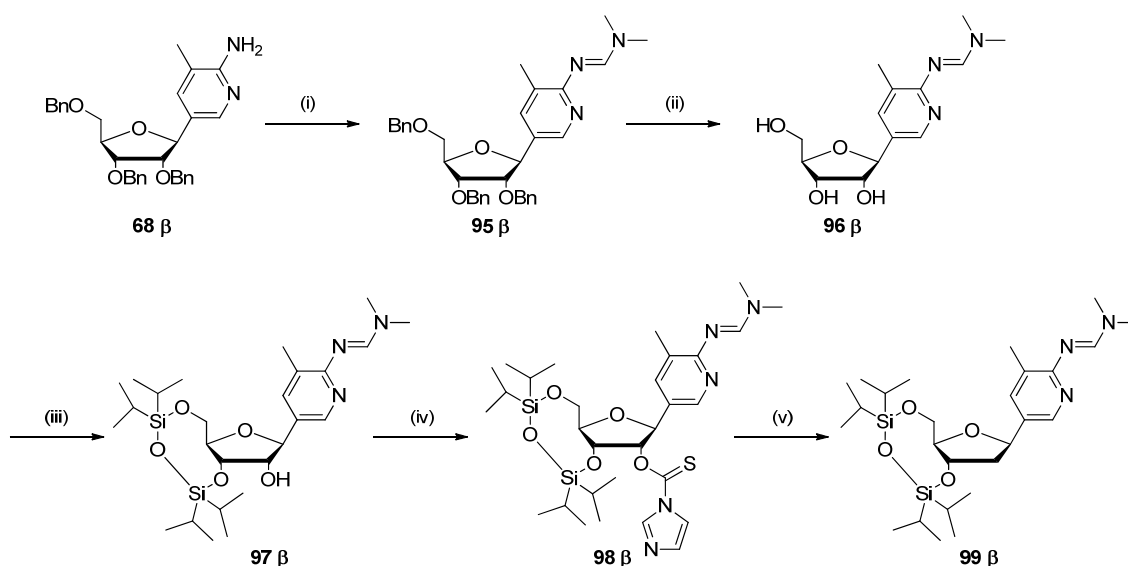


Fig. 2.9: The structure of the free ribo-*C*-nucleoside, **94 β**.

The nucleoside, **68 β** was then protected at the exocyclic amino function using an excess of *N,N*-dimethylformamide dimethylacetal at elevated temperatures to afford **95 β** in *ca.* 68 % yield²⁵⁵ (*scheme. 2.23*). Formamidine protection was introduced as it can be easily cleaved under mildly basic conditions (dilute ammonia²⁵⁶), hence ideal for when incorporating the final monomer into an oligonucleotide. **95 β** was then subjected to Lewis acid conditions of BBr₃ at low temperatures,¹⁵⁸ keeping the formamidine protection intact, to give the *N,N*-protected ribo-*C*-nucleoside, **96 β** in quantitative yield. The free nucleoside, **94 β** obtained during attempt 1, was protected at the exocyclic amino function with formamidine protection which also gave **96 β** in *ca.* 77 % yield.²⁵⁵ Protection of the 3'- and 5'-hydroxyl groups of **96 β** with 1,1,3,3-tetraisopropylidisiloxane in pyridine afforded **97 β** in *ca.* 57 % yield. Activation of the 2'- position of **97 β** with thiocarbonyl diimidazole in *ca.* 55 % yield,¹⁵⁸ followed by

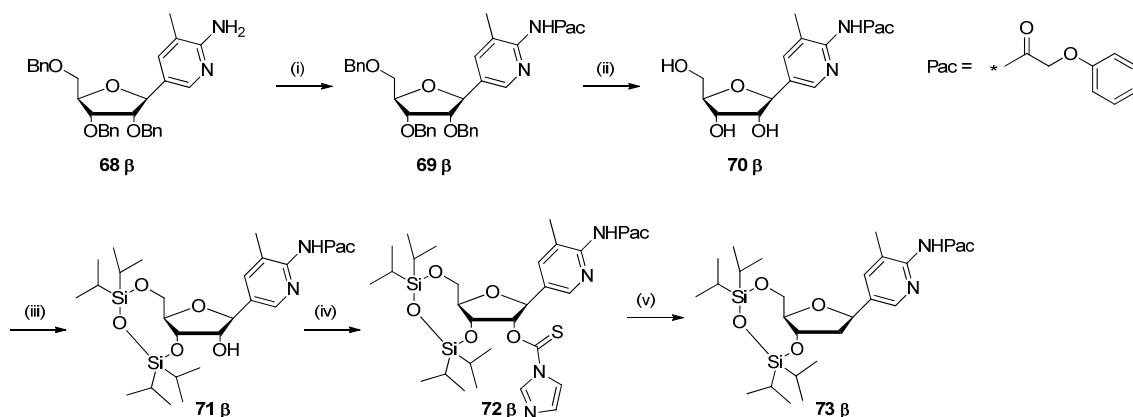
Barton²³¹ homolytic reductive cleavage of the C-O bond provided **99 β**. **99 β** was not isolated as decomposition occurred during its purification by column chromatography, MS analysis on reaction completion confirmed the correct product. After many attempts to improve the yield obtained for the activation of the 2'- position, it remained comparatively low, in comparison to those obtained by Hildbrand *et al.*¹⁵⁸ (74 %) and Dr. Powers²⁰¹ (80 %), when using the Pac protection. It was therefore thought that the dimethylformamidine protection was causing this deficit in yield for this specific reaction. It was decided to go back to the original scheme proposed by Dr. Powers,²⁰¹ and incorporate Pac protection onto the exocyclic amino function, even though the cleavage of this protecting group after oligonucleotide synthesis requires long duration in dilute ammonia.



Scheme. 2.23: Attempt 2 for the synthesis of a ^{Me}P monomer. *Reagents and conditions:* (i) 10 eq N,N-dimethylformamide dimethylacetal, MeOH, 80 °C, 22 hrs, 67.8 %; (ii) 4.5 eq BBr₃ in DCM, -78 °C→r.t, 8 hrs, 99 %; (iii) 1.5 eq (Cl(*i*Pr)₂Si)₂O, pyridine, 0 °C, 1 hr, r.t, 4 hrs, 57 %; (iv) 2.5 eq (Im)₂CS, DMF, r.t, 25 hrs, 55.3 %; (v) 0.2 eq AIBN, 1.15 eq (Me₃Si)₃SiH, toluene, r.t, 1 hr then 80 °C, 5 hrs.^{158,201,255}

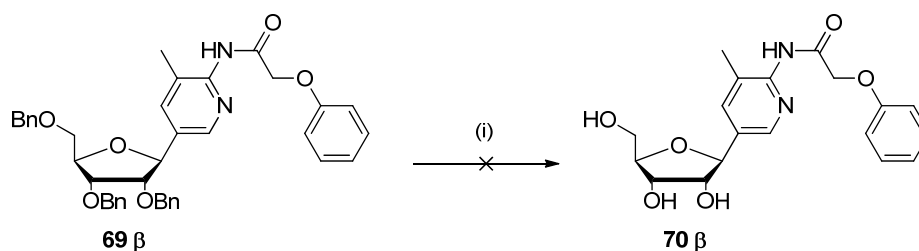
Protection of the exocyclic amino function in **68 β** with phenoxyacetic anhydride afforded **69 β** in *ca.* 80 % yield¹⁵⁸ (scheme. 2.24). Debenzylation under Lewis acid conditions¹⁵⁸ of BBr₃ afforded the *N*-protected ribo-*C*-nucleoside, **70 β** in *ca.* 22 % yield. This low yield was accounted for by the formation of the free ribo-*C*-nucleoside, **94 β** (fig. 2.9) which was isolated in *ca.* 46 % yield; 68 % overall yield for this step. The

phenoxyacetyl (Pac) group is therefore also susceptible to cleavage under the Lewis acid conditions.



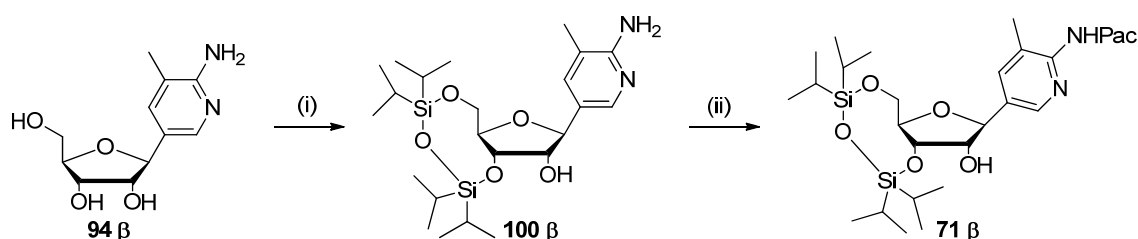
Scheme. 2.24: Attempt 3 for the synthesis of a ^{Me}P monomer. *Reagents and conditions:* (i) 1.7 eq (Pac)₂O, pyridine, r.t, 4 ½ hrs; 79.5 %; (ii) 4.6 eq BBr₃, DCM, -78 °C, 4 hrs, MeOH, -78 °C→r.t, overnight, 68 %; (iii) 1.2 eq (Cl(*i*Pr)₂Si)₂O, pyridine, 0 °C, 1 hr, r.t, 4 hrs, 60.7 %; (iv) 2.5 eq (Im)₂CS, DMF, r.t, 29 hrs, 55.1 %; (v) 2.0 eq AIBN, 1.15 eq (Me₃Si)₃SiH, r.t, 1 hr then 80 °C for 5 ½ hrs, 61.4 %.¹⁵⁸

A popular method to deprotect benzyl ethers is through catalytic hydrogenolysis employing various Pd-catalysts. However, all three attempts outlined in *scheme. 2.25* proved unsuccessful (no reaction had occurred); (i) the first attempt to cleave the benzyl ethers and keeping the Pac group intact was by palladium catalysed hydrogenolysis using H₂, Pd/C and sodium bicarbonate; (ii) the second attempt utilised H₂, Pd/C in ethanol; (iii) the final attempt utilised 50 % aq. acetic acid.



Scheme. 2.25: The unsuccessful attempts for debenzilation of **69 β** keeping the Pac group intact. *Reagents and conditions:* (i) Pd-C, H₂, NaHCO₃, EtOH, r.t, 6 hrs; (ii) Pd-C, H₂ EtOH, r.t, 23 hrs; (iii) Pd-C, H₂, acetic acid solution (50 % in H₂O).^{257,258}

As attempts to improve the debenzylation reaction had proved unsuccessful, both **70 β** and **94 β** were separately protected at the 3'- and 5'-positions, with 1,3-dichloro-1,1,3,3-tetraisopropylidisiloxane¹⁵⁸ to give **71 β** (*ca.* 61 % yield, *see scheme. 2.24*) and **100 β** (*ca.* 50 %, *see scheme. 2.26*) respectively. **100 β** was then protected selectively²⁵⁹ at the exocyclic amino function in the presence of the free 2'-hydroxyl group, with pentafluorophenyl-2-phenoxyethanoate, **101** and provided **71 β** in *ca.* 49 % yield (*scheme. 2.26*).



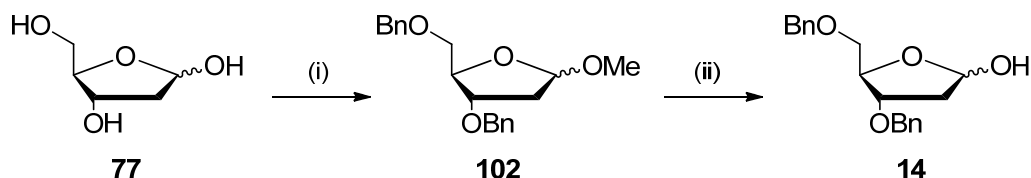
Scheme. 2.26: The conversion of **94 β** to **71 β**. *Reagents and conditions:* (i) 1.2 eq (Cl(*i*Pr)₂Si)₂O, pyridine, 0 °C, 1 ½ hrs, r.t, 4 hrs, 50.3 %; (ii) 1.5 eq Et₃N, 1.2 eq pentafluorophenyl 2-phenoxyethanoate, **101**, CHCl₃, 0 °C, 45 mins, r.t, 21 hrs, 49.2 %.^{158,259}

71 β was then activated at the 2' position with 1,1-thiocarbonyl diimidazole to afford **72 β** in *ca.* 55 % yield which is still comparatively low for reasons unknown. Following Barton²³¹ homolytic reductive cleavage of the C-O bond with AIBN and tris(trimethylsilyl)silane in toluene at elevated temperatures,¹⁵⁸ **73 β** was obtained in *ca.* 61% yield (*scheme. 2.24*). Attempts were made to improve this deoxygenation reaction yield; (i) 1,1'-Azobis(cyclohexanecarbonitrile) has been reported²⁶⁰ to be a more efficient primary radical initiator than AIBN and has been employed to initiate primary radical reactions for C-C bond formation - the attempt that utilised this in the deoxygenation reaction, was unsuccessful; (ii) varying quantities of the AIBN initiator, (0.5-2.0 equivalents); **73 β** was obtained in *ca.* 61 % yield when using 2.0 eq AIBN; 0.5 eq gave an incomplete reaction and yield of *ca.* 66 %; (iii) increasing the amount of tris(trimethylsilyl)silane from 1.15 eq stated in the literature¹⁵⁸ to 1.5 eq resulted in the breakdown of products and hence a yield of *ca.* 37 %. It was concluded that the best combination was 0.5 eq AIBN and 1.15 eq tris(trimethylsilyl)silane, due to recovering the starting material during purification. At this stage there was not enough of **73 β** to proceed through the three remaining steps of the synthetic route to achieve a suitable quantity of the final monomer **76**. The route would have to be repeated from the

beginning on a large scale with the use of the highly expensive TiPDS reagent, which was only required for three steps in this scheme. This led to the development of an alternative, short, concise synthetic route to the ^{Me}P monomer, **76**.

A shorter synthetic route was then constructed, based essentially on the literature by Harusawa *et al.*^{220,221} The route was based on the commercially available starting material, 2'-deoxy-D-ribose.

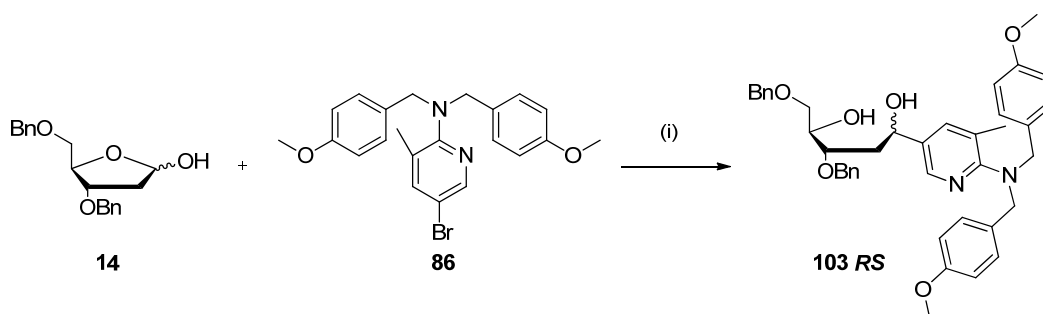
The fully protected 2'-deoxy sugar, **102** was successfully synthesised (*scheme. 2.27*) in *ca.* 95 % yield in a two-step reaction.²⁵³ The 2'-deoxy-D-ribose, **77** was first²⁴⁷ protected as its 1'-*O*-methyl glycoside using acidic methanol, followed by benzylation²⁵³ of the 3'- and 5'-positions using an excess of benzyl bromide and sodium hydride in DMF. The 1'-methyl group was then hydrolysed under acidic conditions to give the perbenzylated 2'-deoxylactol **14**. Initially this acidic hydrolysis was attempted with the method used for the D-ribose derivative,²⁴⁹ 80 % acetic acid solution with concentrated sulphuric acid at 120 °C for 35 mins,²⁰¹ however this resulted in cleavage of the benzyl groups. Instead a milder protocol^{261,262} of solely 80 % acetic acid solution enabled this hydrolysis to proceed in *ca.* 60 % yield.



Scheme. 2.27: Route to the perbenzylated lactol, **14**. *Reagents and conditions:* (i) 0.06 eq CH₃COCl, MeOH, r.t, 1 hr, then 4 eq Benzyl Bromide, 4 eq NaH, DMF, 14 hrs, 95 %; (ii) AcOH (80 % in water), reflux, 120 °C, 2 hrs, 60 %.^{253,261,262}

The reaction of 3',5'-di-*O*-benzyl-2'-deoxy-D-ribose, **14** with excess lithium salt of bis-protected 3-methyl-2-aminopyridine, **86** was initially carried out at the literature reported²²⁰ temperature of – 50 °C, to afford the product **103 RS** in only 18 % yield. In an attempt to improve the yield, the reaction was then attempted at – 78 °C, however, surprisingly no reaction occurred on this occasion. Reese *et al.*²⁰² reported the coupling reaction of the lactol, **78** and the lithium salt of the fully protected base, **79**, at – 78 °C to afford product in 64.4 % yield. The coupling was also attempted *via* the use of a

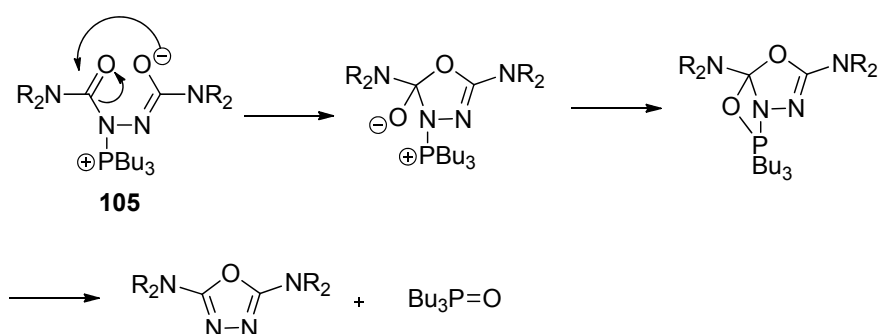
Grignard reagent; magnesium powder was added to a solution of **86** in THF and the reaction mixture was heated to reflux (70 °C). After two hours, the reaction displayed no progression, and the reaction mixture was then cooled to 0 °C for five hours, however, no reaction occurred. The coupling reaction was proving to be sensitive to temperature; finally, the optimum temperature for this reaction was discovered to be - 35 °C. The reaction afforded adduct **103 RS** in *ca.* 92 % yield (*scheme. 2.28*).



Scheme. 2.28: Coupling reaction between the perbenzylated lactol, **14** and the fully protected 2-amino-3-methylpyridine, **86**. *Reagents and conditions:* (i) 1.0 eq *n*BuLi, THF, - 35 °C, 30 mins, 0.4 eq **14** in THF, - 30 °C, 1 hr then r.t, 4 ½ hrs, 92 %.

Initially, Mitsunobu cyclisation of **103 RS** with *N,N,N',N'*-tetramethylazodicarboxamide (TMAD) and tributylphosphine (Bu₃P) in benzene as per Harusawa *et al.*²²⁰ afforded the fully protected *C*-nucleosides **104 α,β** in *ca.* 44.6 %. This mixture was separated into its pure anomers by column chromatography (α:β 1:2). The low yield was thought to be due to a possible competitive side reaction of TMAD taking place²⁶³ (*scheme. 2.29*); where the betaine, **105** produced in the first step of the Mitsunobu reaction cyclises intramolecularly, and hence further reaction cannot occur. Tsunoda *et al.*²⁶⁴ developed 1,6-dimethyl-1,5,7-hexahydro-1,4,6,7-tetrazocin-2,5-dione (DHTD) which is a cyclic Mitsunobu reagent designed to prevent the cyclisation side reaction of acyclic azodicarboxamides. Hence the Mitsunobu cyclisation of **103 RS** was attempted using DHTD and Bu₃P in benzene, to afford **104 α,β** in *ca.* 54.3 % (α:β 2:3). DHTD is not a commercially available reagent however; hence in an attempt to improve the yield further with a commercially available reagent, the cyclisation of the diols was attempted using diisopropylazodicarboxylate (DIAD) and Bu₃P in THF. This initially posed problems in the separation of the anomeric mixture **104 α,β** from the reduced DIAD by-product, as they both had very similar polarities and the reduced form of DIAD is soluble in organic solvents (whereas the reduced TMAD and DHTD was

insoluble in organic solvents and hence easily separated by filtration from the anomeric mixture). Many methods were attempted to remove the reduced DIAD by-product from the anomeric product according to reported procedures²⁶⁵⁻²⁶⁷ either during the work-up or through precipitation using various solvents, however none of these proved successful. Dowex resin 50WX2-200 (H⁺) was then employed as a method to trap the anomeric product to the resin and separate the by-product, once all by-product had been removed, the anomeric product could then be released from the acidic resin with 10 % NH₃ in MeOH: H₂O 1:1.



Scheme. 2.29: Competitive intramolecular side reaction of TMAD during Mitsunobu reactions.²⁶³

This successfully removed the by-product from the anomeric mixture **104 α,β**. However, the acidic nature of the resin was enough to cause partial (affording the anomeric mixture, **106 α,β**), and in some cases, complete cleavage of the PMB groups from the exocyclic amino moiety of the anomers (affording the anomeric mixture, **107 α,β**), resulting in a complex inseparable anomeric mixture. Attempts were then made to separate the anomers by introducing various groups to the exocyclic amino moiety, such as trifluoroacetyl, and dimethylformyl. It was thought that by introducing these protecting groups, they would impose a greater difference in the polarities of the respective α and β anomers. In both cases, separation of the anomers was proven to be impossible. Dowex resin activated with pyridinium ions was then utilised in an attempt to separate the reduced DIAD from the anomeric product **104 α,β** and keeping the PMB groups intact. The resin however was too basic to trap the anomeric product, and hence no separation between product and by-product was achieved. Finally, the anomers were initially separated by column chromatography; this removed some of the reduced DIAD from each anomer. The by-product was then separated from the β anomer of **104**, by submitting the mixture to the next step in the reaction sequence; the removal of the

PMB groups under acidic conditions of trifluoroacetic acid giving **107** β . All traces of the reduced DIAD were successfully removed by column chromatography due to the large difference in the polarities of **107** β and the reduced DIAD. The β -anomer of **107** was obtained in *ca.* 32 % over the two steps (i.e. 32 % for the Mitsunobu cyclisation, on the assumption that the acidic deprotection step was quantitative). It was now possible to obtain the pure anomers with separation from the by-product; however, there was an urgent need to improve the yield. Finally, the Mitsunobu cyclisation of **103** *RS* was achieved using TMAD and Bu₃P in THF to yield **104** α,β in *ca.* 88.2 % yield ($\alpha:\beta$ 2:3)²²⁰ (*scheme. 2.30*).

The configurations about the glycosidic bond were determined with the aid of 2-D ¹H NMR, and NOE studies. It has been reported that the 2'-methylene protons of α - and β -deoxyribonucleosides have been observed to exhibit characteristic patterns in the ¹H NMR.²⁶⁸ The method can be used to determine the anomeric configuration of C-2'-deoxyribonucleosides. The signals for the H^{2'} protons of **104** β spanned a region of 0.38 ppm, which is typical of the β -anomer. The larger region of 0.56 ppm for **104** α provides evidence for the α -configured anomer. Further to this, in the NOESY spectrum of the β -anomer, on the pre-irradiation of the H⁴ proton and one of the two H^{5'} protons separately, a net increase of integrated intensity was experienced by H^{2'} _{β} and not H^{2'} _{α} (*fig. 2.10; see chapter 5, section 5.2 for further details*). In the NOESY spectrum of the α -anomer, on the pre-irradiation of the H⁴ proton, a net increase of integrated intensity was experienced by H^{2'} _{α} and not H^{2'} _{β} ; on the pre-irradiation of the H^{1'} proton, a net increase of integrated intensity was experienced by H^{2'} _{β} and not H^{2'} _{α} . Most importantly, on the pre-irradiation of one of the two H^{5'} protons, a net increase of integrated intensity was experienced by H^{1'}, this provides solid evidence for the configuration about the C^{1'} position in this anomer (*fig. 2.11*). In addition, it is apparent that H^{2'} _{β} resonates at a higher field/lower chemical shift than H^{2'} _{α} in the β -configured anomer, and that H^{2'} _{α} resonates at a higher field/lower chemical shift than H^{2'} _{β} in the α -configured anomer; Reese *et al.*²⁰² have also observed this in their NMR studies of the free nucleoside of MeP (**1**).

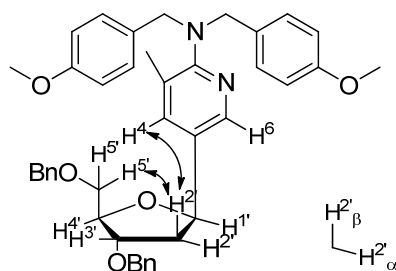


Fig. 2.10: A diagrammatic illustration displaying the through space proton-proton interactions indicated by the NOESY spectra of **104 β** .

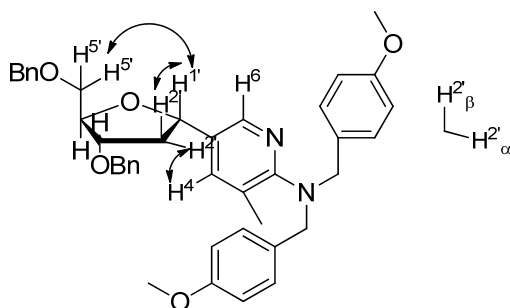
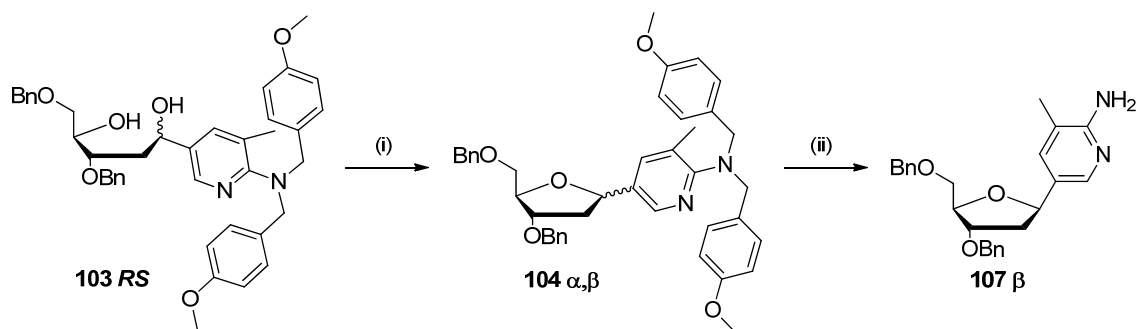


Fig. 2.11: A diagrammatic illustration displaying through space proton-proton interactions indicated by the NOESY spectra of **104 α** .

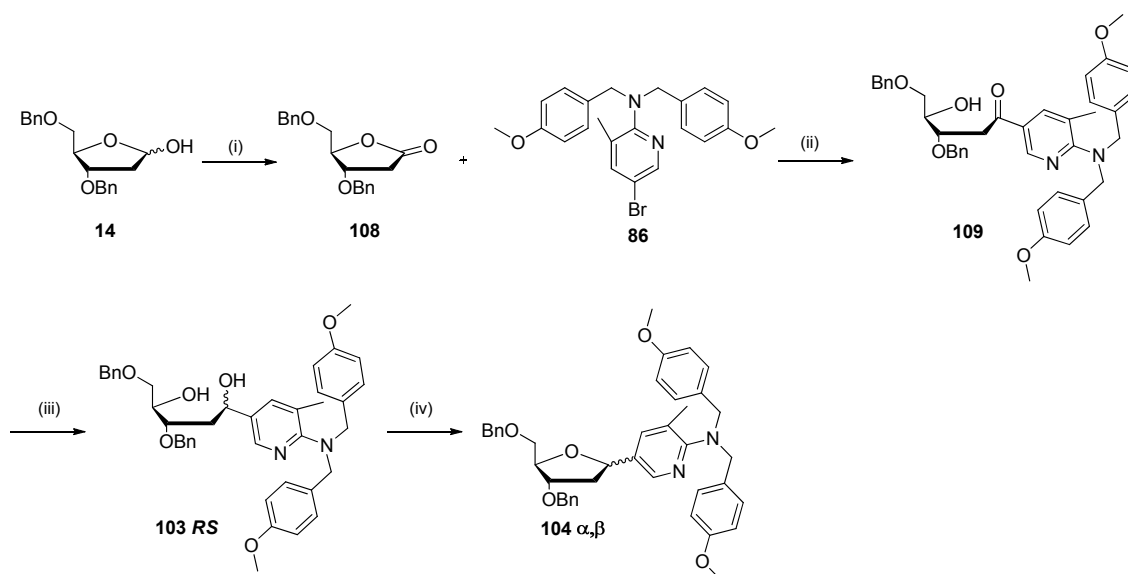
Cleavage of the PMB groups in **104 β** in trifluoroacetic acid: DCM 10:7 afforded **107 β** in *ca.* 98.9 % yield.²⁰² Starting from **103 *RS***, **107 β** was obtained in *ca.* 52.3 % over the two steps. When compared to the previous method using DIAD, TMAD gives the optimum yield and easier separation of the reduced azo-reagent, hence the method of choice for this step.



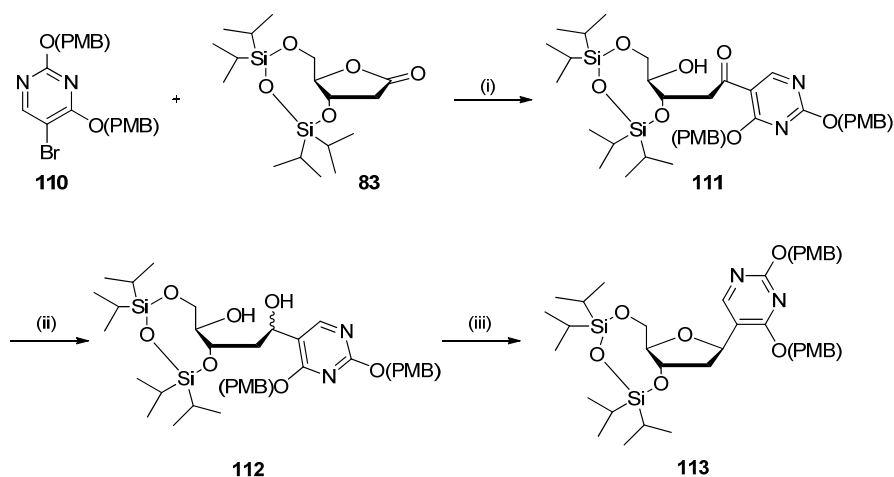
Scheme. 2.30: Mitsunobu cyclisation of diols **103 RS** followed by cleavage of the PMB groups in **104 α,β**. *Reagents and conditions:* (i) 1.3 eq Bu₃P, 1.3 eq TMAD, Benzene, 0 °C→r.t, 22 hrs, 44.6 %; (i) 1.5 eq Bu₃P, 1.5 eq DHTD, benzene, 0 °C→r.t, 25 hrs, 54.3 %; (i) 1.6 eq Bu₃P, 1.6 eq DIAD, THF, 0 °C→r.t, 5 ½ hrs, 32 %; (i) 1.6 eq Bu₃P, 1.6 eq TMAD, THF, 0 °C→r.t, 6 ¼ hrs, 88.2 %; (ii) Trifluoroacetic acid: DCM 10:7, r.t, 5 ½ hrs, 98.9 %.^{202,220,264}

The anomeric ratio obtained after Mitsunobu cyclisation was $\alpha:\beta$ 2:3, whereas in the case of Harusawa *et al.*²²⁰ they obtained an almost exclusive β -configured mixture (discussed in more depth in section 2.2.3.3.1) with ratio of $\alpha:\beta$ 1:5 from a 1:1 mixture of the *R* and *S* epimers. They gave reasoning for this stereocontrolled synthesis of the *C*-nucleoside to the extra stabilisation from the hydrogen bond formed between the nitrogen in the imidazole and the 5'-OH group in the sugar (see *fig. 2.4*) in the intermediate. They also encountered exclusively the β -configured nucleoside of the ribose derivative from a 1:1 mixture of the epimeric diols. This excellent β -glycosylation was accounted for by the stabilisation from the hydrogen bonding described for the 2'-deoxy version, but also due to the presence of a 2'-*O*-benzyl substituent. The 2'-*O*-benzyl substituent creates unfavourable sterics and electronic repulsion from orbital overlap in the intermediate, when the orientation about the C-1' is in the *S*-configuration. Yokoyama *et al.*²⁶⁹ however reported the synthesis of *C*-nucleosides having typical aromatic heterocycles as the base moiety, in which the coupling reaction of the lithiated salt with the ribofuranose derivative is stereoselective depending on the heteroatom within the heterocyclic base. The subsequent Mitsunobu cyclisation is then stereospecific; orientation of the glycosidic linkage is controlled by the C-1' configuration of the substrate: one isomer affords the α -anomer and the other isomer the β -anomer. It therefore, may have been useful in the case studied here to separate the epimeric mixture of diols, **103 RS**, and then proceed with Mitsunobu cyclisation separately to observe which epimer produced the α -anomer and which the β -

anomer. This would significantly simplify the Mitsunobu cyclisation when using DIAD, and the removal of the reduced DIAD by-product by Dowex in the H⁺ form would be successful. The mixture of cleavage products obtained would no longer be a problem as they would all have the same configuration about the C-1'. This may have also improved the yield of pure **104** β isolated. The product, **104** β isolated after Mitsunobu cyclisation using TMAD was separated by column chromatography from its α -anomer. Although separable, the anomers have very similar polarities within the system used for separation, hence resulting in some portion of the purified β -product being contaminated with its α -anomer. This portion could not be used for successive steps of the scheme, therefore lowering the overall quantity of product obtained at the end of the synthesis. Alternatively, as mentioned in *section 2.3.1*, as Reese *et al.*²⁰² suggested; the β -stereoselectivity in this case may be increased if the synthesis was started with the perbenzylated 2'-deoxy lactone **108** (*scheme. 2.31*) and then reducing the putative intermediate hydroxy-ketone, **109** with L-Selectride before carrying out the Mitsunobu cyclisation reaction. Reese *et al.*²⁰² used this method in the synthesis of 2'-deoxypseudouridine (*scheme. 2.32*). The protected lithiopyrimidine of **110** was reacted with the corresponding lactone, **83** at -100 °C to give the hydroxy-ketone, **111**. This was then reduced with L-selectride, the diol mixture, **112** was then cyclised under Mitsunobu conditions to give almost entirely the β -anomer of the fully protected 2'-deoxypseudouridine, **113**. They suggested that if it is assumed that the Mitsunobu reaction was highly stereospecific, the L-selectride reduction was therefore highly stereoselective.



Scheme. 2.31: A potential route to improve the β -glycosylation of the Mitsunobu cyclisation reaction by starting the synthetic route with the lactone **108**. *Reagents and conditions:* (i) 5 mol % TPAP, 1.6 eq NMO, 4 Å powdered molecular sieves, DCM, 0 °C \rightarrow r.t.; (ii) 1.03 eq *n*BuLi, THF, -78 °C, then 1.07 eq **108**, THF, -78 °C \rightarrow 0 °C; (iii) L-Selectride, THF, -78 °C \rightarrow r.t.; (iv) 1.6 eq Bu₃P, 1.6 eq TMAD, THF, 0 °C \rightarrow r.t.; (iv) 1.6 eq Bu₃P, 1.6 eq DIAD, THF, 0 °C \rightarrow r.t.^{202,220}



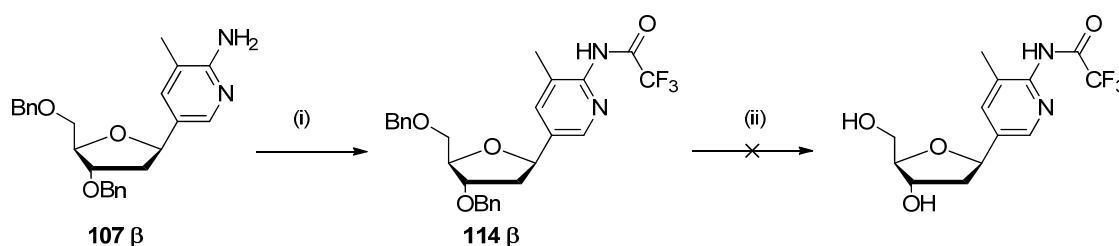
Scheme. 2.32: Synthesis of 2'-deoxypseudouridine, **113** by Reese *et al.*²⁰² *Reagents and conditions:* (i) 1.05 eq *n*BuLi, THF, hexane, -100 °C, 30 mins, then 0.8 eq lactone, **83**, THF, -100 °C \rightarrow -78 °C, 3 hrs; (ii) 1.2 eq L-Selectride, -78 °C \rightarrow r.t.; (iii) 1.2 eq DIAD, 1.2 eq Ph₃P, THF, r.t, 3 hrs, 43.6 % over steps (i)-(iii).

Alternatively, the use of bulky protecting groups for the 3'- and 5'-hydroxyl groups such as *tert*-butyldiphenylsilyl (TBDPS) may influence the stereoselectivity in the coupling reaction of the lithiated base, **86** with the lactol, **14**. Hence, producing a greater

epimeric ratio bias towards the diol that cyclises under Mitsunobu conditions to give the β -anomer.

The suggestions made above may be worth attempting in order to optimise the Mitsunobu cyclisation to give better β -glycosylation of compound **104**, and hence improve the overall yield of the ^{Me}P monomer.

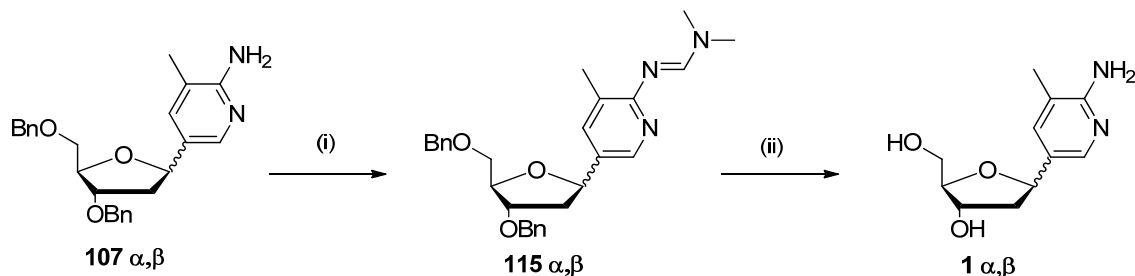
The exocyclic amino moiety in **107** β was then protected with trifluoroacetic anhydride in pyridine at 0 °C to afford **114** β in *ca.* 70 % (*scheme. 2.33*).^{251,254} Debenzylation by hydrogenolysis was then attempted by method of transfer hydrogen²⁵⁴ using 20 % Pd(OH)₂-C with cyclohexene in ethanol. This reaction unfortunately caused cleavage of the trifluoroacetyl group to afford **107** β . Trifluoroacetamide is an easily cleavable amide, previous citations²⁷⁰ have reported solvolysis in methanol or ethanol at ambient temperature with the by-product being the corresponding alkyl trifluoroacetate. Ethyl acetate was then employed as the reaction solvent; however the reaction proceeded very slowly only to cleave the trifluoroacetyl group. The trifluoroacetyl group was proving to be too labile to be used for protection of the exocyclic amino function, hence was discontinued for the use as a protecting group. It was also concluded that cyclohexene was not a sufficient source of hydrogen to allow hydrogenolysis to occur.



Scheme. 2.33: *Reagents and conditions:* (i) 1.5 eq trifluoroacetic anhydride, pyridine, 0 °C, r.t, 2 ½ hrs, 70 %; (ii) 20 % Pd(OH)₂-C, 1 eq cyclohexene, ethanol, 80 °C, 6 hrs; (ii) 20 % Pd(OH)₂-C, 1 eq cyclohexene, ethyl acetate, 80 °C.^{251,254}

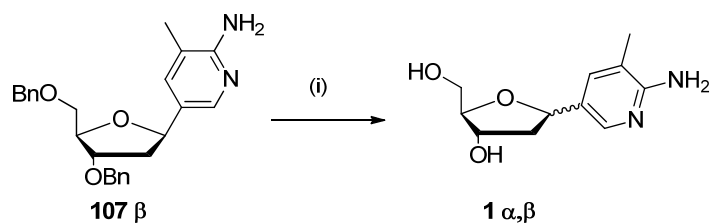
In an examination to investigate if dimethylformyl could withstand the conditions of debenzylolation, dimethylformamidine protection was introduced to the exocyclic amino moiety of **107** α,β using *N,N*-dimethylformamide dimethylacetal in methanol at elevated temperature²⁵⁵ to give **115** α,β . This then underwent hydrogenolysis by transfer hydrogen using 20 % Pd(OH)₂-C with formic acid in methanol,²⁷¹ although

debenzylation had occurred as proven by NMR, the dimethylformamidine protection had also been cleaved, affording **1** α,β (*scheme. 2.34*). The protecting group was thought to have been cleaved during neutralisation using NaOH.



Scheme. 2.34: *Reagents and conditions:* (i) 10 eq *N,N*-dimethylformamide dimethylacetal, MeOH, 70 °C, 26 hrs; (ii) 20 % Pd(OH)₂-C, formic acid, MeOH, 50 °C, 6 hrs.^{255,271}

Various attempts for the successful cleavage of the benzyl groups in **107** were carried out. Using boron tribromide on each of the anomers, **107** α and **107** β separately resulted in complete debenzylation (*scheme. 2.35*). However, in each case the presence of an anomeric mixture, **1** α,β was confirmed by NMR results, indicating ring opening and re-closing had occurred. The strong acidity of HBr formed during the reaction was thought to be the major cause of this ring-opening, however, it was also thought that extra stabilisation within the structure of the molecule was enabling this unfavourable ring opening reaction to proceed. The extra stabilisation was thought to come from either; (i) the lone pairs on the exocyclic amino function being able to conjugate through the nucleobase part of the nucleoside and hence acting essentially as a benzyl group (*fig. 2.12*) or; (ii) the 2'-protons are acidic enough to be taken up by a bromide from the BBr₃ Lewis acid (*fig. 2.13*). As a result of ring-opening in both cases, and the presence of the double bond at the C-1' position, the C-4' oxo-anion can attack from either face of the double bond (*path a gives α , while path b gives β*) producing an anomeric mixture. Interestingly in both cases, when starting from either α or β , the anomeric ratio is in favour of the β -anomer. Lewis acid mediated debenzylation using BBr₃ is therefore not a suitable method to be used for the β -anomer. However, it could be a potential method to increase the β -product when starting with **107** α , providing a suitable separation technique is developed to isolate the β -anomer from the anomeric mixture.



Scheme. 2.35: Debenzylation of **107 β** using BBr_3 causing ring-opening and re-closing forming **1 α, β** .
Reagents and conditions: (i) 3.5 eq BBr_3 , DCM, -78°C , 1 $\frac{1}{2}$ hrs, MeOH, $-60^\circ\text{C} \rightarrow \text{r.t.}$

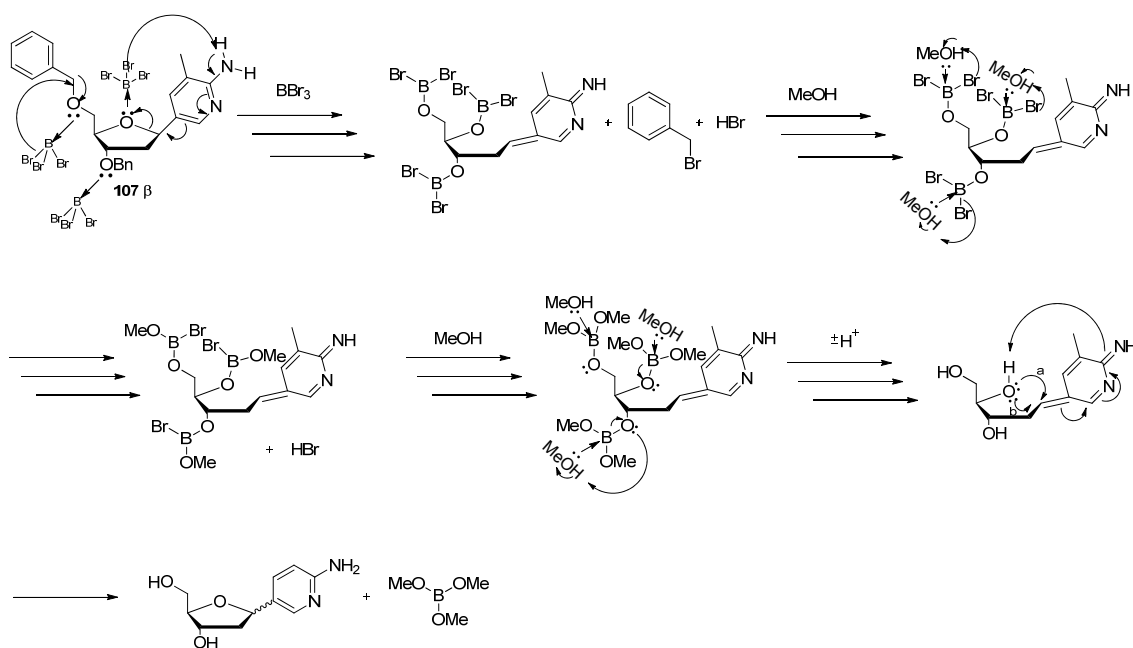


Fig. 2.12: Mechanism 1 for the ring opening of the 2'-deoxy ribose sugar, where the base acts as a benzyl group as a result of stabilisation through lone pair conjugation from the exocyclic amino moiety.

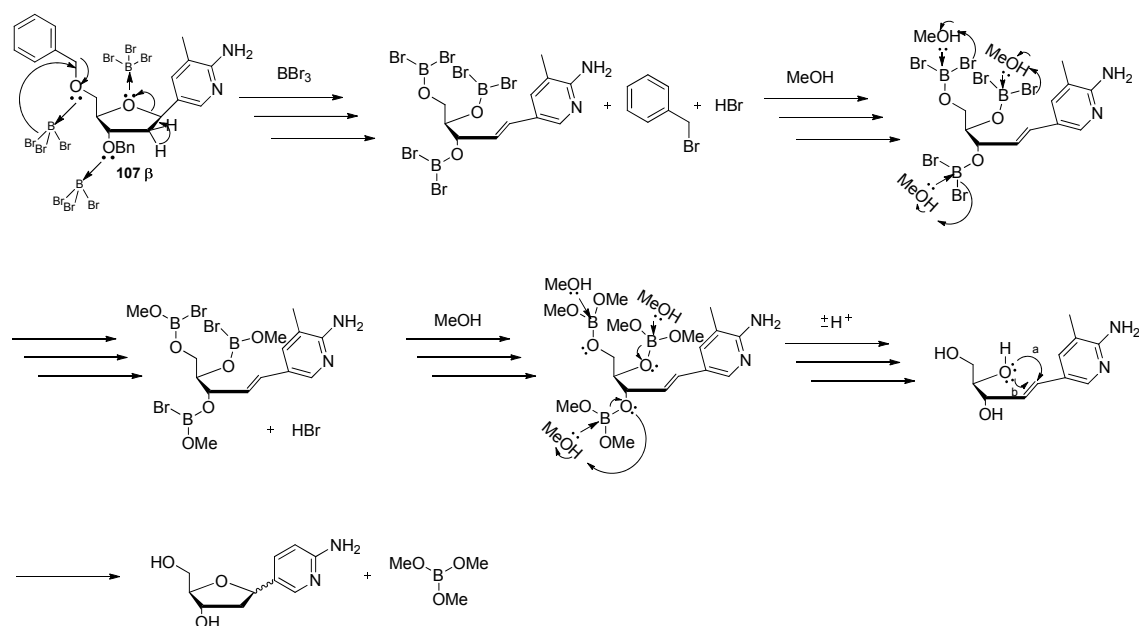
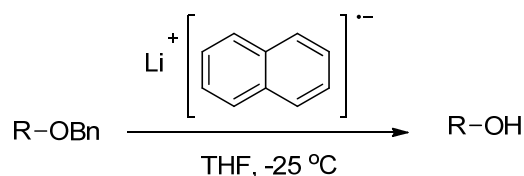


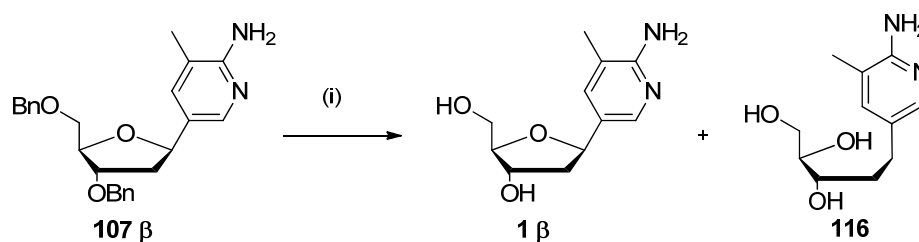
Fig. 2.13: Mechanism 2 for the ring opening of the 2'-deoxy ribose sugar, where the acidic 2'-protons are taken up by a bromide from the BBr_3 Lewis acid.

A convenient method for debenzylation reported in the literatures²⁷²⁻²⁷⁴ uses naphthalene-catalysed lithiation (*scheme. 2.36*), which is a mild, highly effective method for the reduction of benzyl ethers to the corresponding alcohols. This method was attempted for the removal of the benzyl groups from **107 β**, however proved to be unsuccessful, with no reaction occurring.



Scheme. 2.36: Reductive debenzylation with Lithium Naphthalenide.²⁷²

Debenzylation of **107 β** was found to occur under hydrogenolysis conditions, using $\text{Pd}(\text{OH})_2\text{-C}$ as the catalyst, H_2 gas, in ethanol, at elevated temperatures (*scheme. 2.37*). Reaction duration was very slow however (3 days for completion), and afforded **1 β** in *ca.* 25 %. The low yield was accounted for by the formation of an open-chain product, **116**, which was the major product, it was formed by a competitive side reaction (*see fig. 2.14*). The structure (*scheme. 2.37*) was confirmed by MS and NMR analyses.



Scheme. 2.37: Debenzylation of **107 β** using palladium catalysed hydrogenation, to give the desired product **1 β** and the open-chain side product, **116**. *Reagents and conditions:* (i) Pd(OH)₂-C, EtOH, H₂, 60 °C, 80 hrs, 25.1 %.

It was thought that the ring oxygen becomes hydrogenated, along with hydrogenolysis of the 3'- and 5'- benzyl ethers, due to the 3-methyl-2-aminopyridine nucleobase acting as another benzyl group (*fig. 2.14*), as was thought a possible mechanism when using BBr₃.

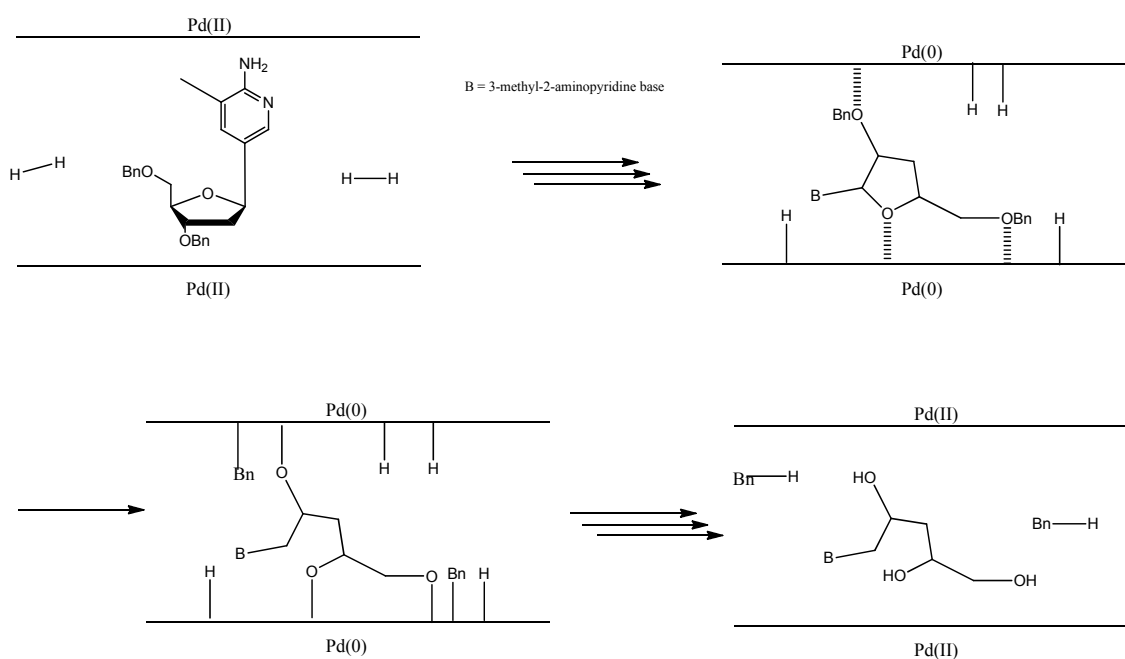


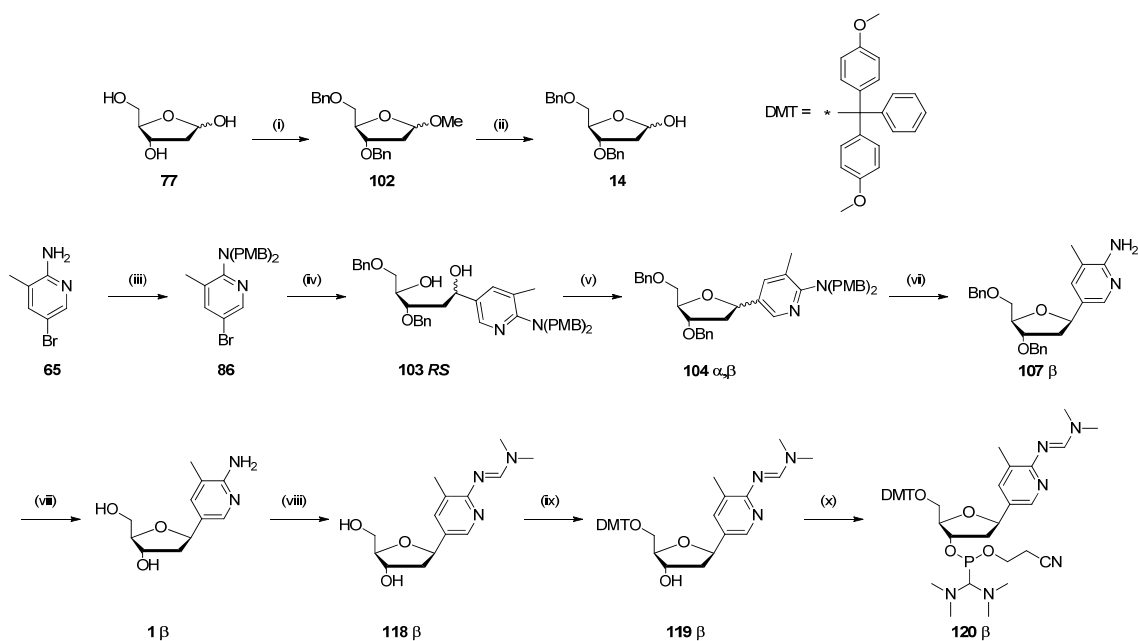
Fig. 2.14: Possible mechanism for the opening of the sugar ring in **107 β** to form the open-chain side product **116**.

Using MS and NMR analyses, it was observed that the product, **1 β** and intermediate, **117 β** (5'-OH, 3'-OBn) formed before the open-chain side product. On stopping the hydrogenolysis reaction once **107 β** had completely disappeared, the product and intermediate could be isolated. The intermediate (**117 β**) was then put under palladium

catalysed hydrogenolysis to allow for product formation. In doing this the overall yield to form **1** β was increased to *ca.* 31 %. This method however, was very time consuming; complete disappearance of **107** β was observed after > 21 hrs and that of **117** β was observed after 31 hrs.

It was concluded at this stage that the absence of a 2'-modification on the sugar ring causes problems under the conditions used for the cleavage of the benzyl groups. Hydrogenolysis conditions however are the mildest; hence the debenzylation step was not optimised further at this stage.

Protection of the exocyclic amino function in **1** β with *N,N*-dimethylformamide dimethylacetal in methanol at elevated temperatures gave **118** β in *ca.* 61 % yield.²⁵⁵ The 5'-position was then protected with 4,4'-dimethoxytrityl chloride in pyridine to afford **119** β in *ca.* 62.4 % yield¹⁵⁸. **119** β was then subjected to conventional phosphitylation conditions of 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite and *N,N*-diisopropylethylamine (Hünig's base) in DCM affording **120** β in *ca.* 79 % yield²⁷⁵ (*scheme 2.38*).



Scheme. 2.38: The synthetic route to ^{Me}P monomer, **120 β**. *Reagents and conditions:* (i) 0.06 eq CH₃COCl, MeOH, r.t, 1 hr, then 4 eq Benzyl Bromide, 4 eq NaH, DMF, 14 hrs, 95 %; (ii) AcOH (80 % in water), reflux, 120 °C, 2 hrs, 60 %; (iii) 2.5 eq 4-methoxybenzyl chloride, 2.5 eq NaH, DMF, r.t, 4 hrs 35 mins, 76.4 %. (iv) 1.0 eq *n*BuLi, THF, - 35 °C, 30 mins, 0.4 eq **14** in THF, - 30 °C, 1 hr then r.t, 4 ½ hrs, 92 %; (v) 1.6 eq Bu₃P, 1.6 eq TMAD, THF, 0 °C→r.t, 6 ¼ hrs, 88.2 %; (vi) Trifluoroacetic acid: DCM 10:7, r.t, 5 ½ hrs, 98.9 %; (vii) Pd(OH)₂-C, EtOH, H₂, 60 °C, 80 hrs, 31 %; (viii) Excess *N,N*-dimethylformamide dimethylacetal, MeOH, 70 °C, 20 hrs, 61 %; (ix) 1.26 eq DMTCl, pyridine, r.t, 2 hrs 20 mins, 62.4%; (x) 2 eq DIPEA, 1.3 eq 2-cyanoethyl-*N,N*-diisopropylchloro phosphoramidite, DCM, r.t, 1 ¾ hrs, 79.1 %.^{158,202,220,253,255,261,262,275}

To summarise, the synthetic route developed here (*scheme. 2.38*) essentially followed the works reported by Harusawa *et al.*²²⁰, Hildbrand *et al.*¹⁵⁸, Reese *et al.*²⁰² and the published thesis of Dr Vicki Powers²⁰¹; a 10-step scheme towards ^{Me}P monomer, **120 β** with overall yield of *ca.* 4.5 % starting from the coupling reaction (*step (iv) in scheme. 2.38*). The free nucleoside, ^{Me}P, **1 β** can be obtained in 7 steps with overall yield of *ca.* 14.9 %. The overall yield to obtain the perbenzylated lactol, **14** was *ca.* 57 % and that to obtain the fully protected base, **86** was *ca.* 76.4 %. The step that pulled the overall yield to **120 β** down was the low yielding debenzylation (*step (vii) in scheme. 2.38*), and this was accounted for by the formation of the open-chain side product, **116**. It can be seen that when the overall yield to the free nucleoside, ^{Me}P (14.9 %) obtained here is compared to that obtained by Hildbrand *et al.*¹⁵⁸ (10.1 %) and Reese *et al.*²⁰² (7.8 %), it is the best route. The step that requires urgent optimisation is *step (vii) in scheme. 2.38*;

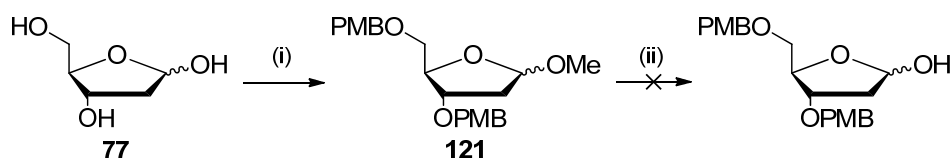
the debenzylation. If this was optimised, the route would be an efficient, concise synthetic procedure to a very useful *C*-nucleoside and Cytidine-analogue.

This *C*-nucleoside derivative was incorporated as protonated Cytidine equivalents in a homopyrimidine oligonucleotide. Footprinting studies indicated that the dimethylformamidine protecting group used for the exocyclic amino moiety had not undergone complete cleavage under standard oligonucleotide deprotection conditions (concentrated aqueous ammonia at room temperature for 24 hrs) (*see chapter 3 for further details*). Cleavage conditions for **118** β had been determined prior to this and the dimethylformamidine protection was shown to be completely removed in 90 % aqueous ammonia in ethanol at 55 °C for 10 hrs. Normally the conditions required for cleavage become milder when the nucleoside is incorporated within the oligo. However, this was not the case as proven by the footprinting results. Therefore, this protecting group was proving unsuitable to be used within oligos, especially those containing psoralen (*discussed further in chapter 3*); the search for a new exocyclic amino protecting group commenced.

Initially, the aim was to provide the ^{Me}P monomer with trifluoroacetyl protection at the exocyclic position; although capping could not be used during oligonucleotide synthesis, the protecting group is very easily cleaved, as mentioned previously. The debenzylation step was also in urgent need to be optimised; a more readily cleavable protecting group for the 3'- and 5'-hydroxyl group such as paramethoxy benzyl, para-nitrobenzyl, 2,4-dinitrobenzyl, 2,4-dimethoxybenzyl, benzhydryl etc may be necessary. The α -anomer obtained from *scheme. 2.38* after *step (v)* was used for the initial tests and optimisations for the new scheme. This was to ensure that the β -anomer was not wasted on trial attempts.

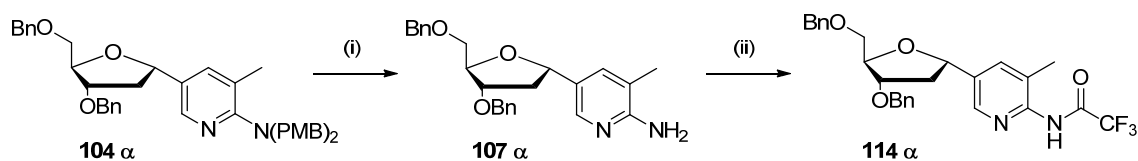
Paramethoxy benzyl protection could be used in place of the benzyl protection for the 3'- and 5'- positions of the 2'-deoxyribose. Hence post Mitsunobu cyclisation, all protecting groups could potentially be removed in one step using trifluoroacetic acid in DCM. The 2'-deoxy-D-ribose, **77** was first protected as its 1'-*O*-methyl glycoside using acidic methanol, followed by paramethoxybenzylation of the 3'- and 5'-positions using an excess of benzyl bromide and sodium hydride in DMF and afforded **121** in *ca.* 98 % yield^{202,247,253} (*scheme. 2.39*). Hydrolysis of the 1'-methyl group was then attempted

under acidic conditions of 80 % acetic acid solution,^{261,262} the reaction was thought to have unfortunately caused partial cleavage of the paramethoxybenzyl groups. Hence, the use of this group as protection for the 3', and 5' positions of the sugar ring was discontinued.



Scheme. 2.39: *Reagents and conditions:* (i) 0.06 eq CH₃COCl, MeOH, r.t, 3 hrs 20 mins, then 4 eq *p*-methoxybenzyl chloride, 4 eq NaH, DMF, 15 hrs, 98 %; (ii) AcOH (80 % in water), 110 °C, 1 ½ hrs.^{202,247,253,261,262}

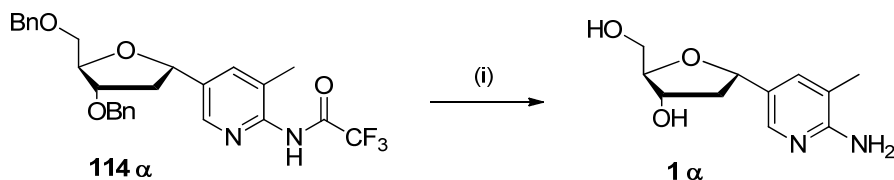
Optimisation of the debenzylation step then commenced. Successful removal of the PMB groups within **104 α** under acidic conditions of trifluoroacetic acid in DCM afforded **107 α** in *ca.* 89 % yield²⁰² (*scheme. 2.40*). It was thought that the addition of an electron-withdrawing group such as TFA to the exocyclic amino moiety in **107 α** would prevent the 2-amino-3-methylpyridine ring acting as a benzyl group, and hence prevent sugar ring opening during conditions used for debenzylation. Trifluoroacetylation of the exocyclic amino function in **107 α** was attempted using excess ethyl trifluoroacetate in methanol at 0 °C²⁷⁶ which proved unsuccessful, this method was repeated in DMF however no progression was observed. Finally, trifluoroacetylation was successful using trifluoroacetic anhydride in pyridine²⁵¹ affording **114 α** in *ca.* 97 % yield.



Scheme. 2.40: *Reagents and conditions:* (i) Trifluoroacetic acid: DCM 1:1, r.t, 4 ½ hrs, 89 %; (ii) 1.5 eq trifluoroacetic anhydride, pyridine, 0 °C → r.t, 2 ¾ hrs, 97 %.^{202,251}

Debenzylation by hydrogenolysis was attempted with hydrogen gas, Pd(OH)₂-C catalyst in ethanol at elevated temperatures (60 °C). However, after a total of 82 hrs reaction there was still starting material present, and it was thought to contain breakdown products. Successful cleavage of the benzyl groups was then achieved using boron

trichloride in DCM at $-78\text{ }^{\circ}\text{C}$ ²⁷⁷ (scheme. 2.41), ring opening had not occurred as was proven by NMR, however the TFA group had been cleaved. **1 α** was afforded in *ca.* 96 % yield.



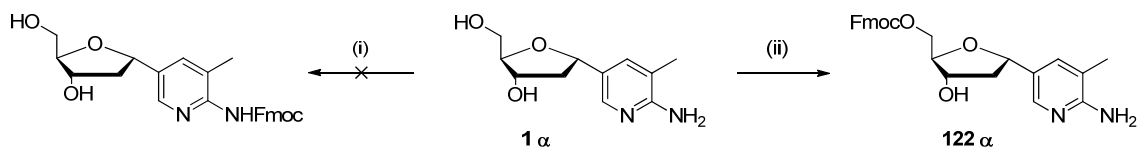
Scheme. 2.41: Reagents and conditions: (i) 3 eq BCl_3 , DCM, $-78\text{ }^{\circ}\text{C}$, 7 hrs, then MeOH, $4\text{ }^{\circ}\text{C}$, 18 hrs, 96 %.²⁷⁷

It is important to note that care had to be taken with this reaction, as prolonged reaction time was observed to cause ring-opening and re-closing forming an anomeric mixture. Also, the methanol added to quench the reaction mixture had to be in excess in order to efficiently terminate the reaction. It was observed that after 72 hrs, starting with **114 α** the reaction gave **1 α,β** in a 1:1 anomeric mixture in *ca.* 72 % yield. This would be useful in order to improve the yield of β -anomer, starting from the α -configured starting material, as mentioned previously when using boron tribromide, and needs to be investigated further.

Attempts to protect the exocyclic amino moiety post debenzylation, with TFA using: (i) trifluoroacetic anhydride in pyridine²⁵¹; (ii) trifluoroacetic anhydride in the presence of potassium carbonate in dry 1,4-dioxane at $25\text{ }^{\circ}\text{C}$ ²⁷⁸; (iii) ethyl trifluoroacetate, DMAP in pyridine, proved unsuccessful; the group was observed to be too labile as reported by Bellamy *et al.*²⁷⁰ Attention turned to the use of Fmoc (Fluorenylmethyloxycarbonyl) as a protecting group; it is readily cleaved under mild conditions with base (10 % piperidine in DMF).²⁵⁴ The Fmoc group would therefore be appropriate for use in the end monomer, which will be incorporated into oligos and undergo standard oligonucleotide deprotection conditions.

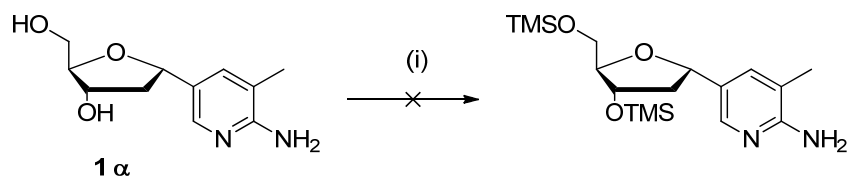
Initial attempts to selectively protect the exocyclic amino moiety of **1 α** using N-(9-Fluorenylmethoxycarbonyloxy) succinimide (Fmoc-OSu)^{279,280} were not successful (scheme. 2.42). When the reaction was carried out with the more reactive Fmoc-Cl in pyridine, **122 α** was isolated; the 5'-position of the sugar ring had become protected

(*scheme. 2.42*). Hence indicating that the 5'-position required protection prior to Fmoc protection of the exocyclic amino position.



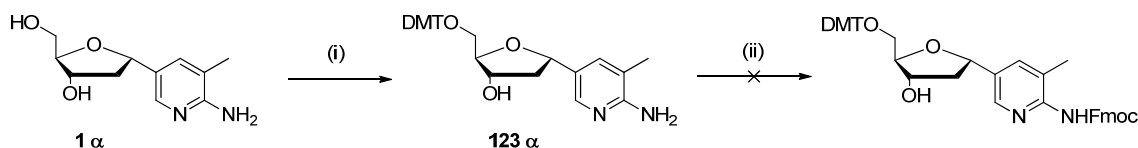
Scheme. 2.42: The failed attempts to protect the exocyclic amino moiety in **1** α with Fmoc protecting group. *Reagents and conditions:* (i) 2 eq NaHCO_3 , 1.1 eq Fmoc-OSu, 1,4-dioxane, 3 °C, 1 hr 50 mins, then r.t, 3 hrs, then 50 °C, 16 hrs; (i) 1.1eq, DIPEA, 1 eq Fmoc-OSu, pyridine, r.t, 1 hr 20 mins, then 55 °C, 16 hrs; (ii) 5 eq Fmoc-Cl, pyridine, r.t, 7 hrs.^{279,280}

Protection of the 3'- and 5'-positions of the sugar ring in **1** α was attempted using trimethylsilyl chloride (TMS-Cl) in pyridine; however no reaction occurred on this occasion (*scheme. 2.43*).



Scheme. 2.43: The failed attempt to protect the 3' and 5' positions in **1** α . *Reagents and conditions:* (i) 15 eq TMS-Cl, pyridine, r.t, 7 hrs, then 45 °C, 17 hrs.

The 5'-position was then tritylated with DMT-Cl in pyridine¹⁵⁸ to afford **123** α in a poor yield of *ca.* 23 %. Attempts were then made to protect the exocyclic amino moiety using Fmoc-Cl with DMAP in pyridine at room temperature²⁸¹, however these were unsuccessful (*scheme. 2.44*).

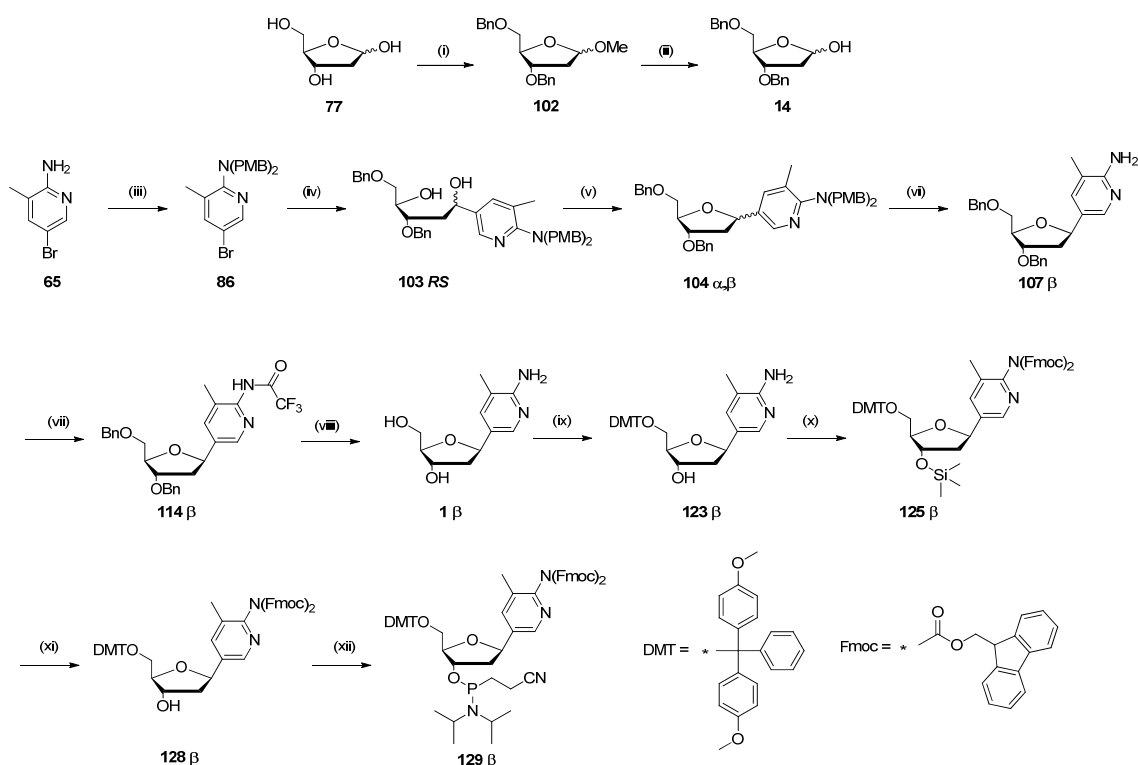


Scheme. 2.44: (i) 1.26 eq DMT-Cl, pyridine, r.t, 3 hrs, 23 %; (ii) Excess Fmoc-Cl, pyridine, r.t, 24 hrs.^{158,281}

It was apparent that in order to protect the exocyclic amino function, the 3'-position of the sugar ring required protection; this was successful using TMS-Cl in pyridine. Immediate Fmoc protection of the exocyclic amino moiety could then proceed using an excess of Fmoc-Cl, producing a mixture of mono- and bis-Fmoc protected nucleosides, **124 α** and **125 α** respectively in *ca.* 49 % yield, with the majority as **125 α**; **125 α** was the product required (*see scheme. 2.45 for the structures of the β equivalents*). Many attempts were made to cleave the TMS group from the 3'-position; (i) pyridine, water, r.t; (ii) sat. NaHCO₃, pyridine; (iii) MeOH, pyridine, K₂CO₃, 0 °C. Cleavage was not successful using the first 2 methods, cleavage occurred under conditions of (iii) however due to the potassium carbonate being highly basic, this resulted in additionally cleaving the Fmoc group. This is as far as optimisation went using the α-anomer.

Successful cleavage of the 3'- and 5'-benzyl groups in **114 β** occurred without disruption of the sugar ring, using boron trichloride in DCM at -78 °C,²⁷⁷ to afford **1 β** in quantitative yield. Tritylation of the 5'-position of the sugar ring went in *ca.* 30 % yield¹⁵⁸ affording **123 β**. This low yield was partly accounted for by the formation of bis DMT protected nucleoside, **126 β** in *ca.* 10 %, due to a slight excess of reagent used. However, this overall yield of *ca.* 40 % is still relatively low for the tritylation reaction, and may be due to the exocyclic amino moiety being unprotected. The 3'-position of the sugar was protected with TMS using TMS-Cl in pyridine. The exocyclic amino moiety was then protected using Fmoc-Cl in pyridine, producing a mixture of products; mono-Fmoc protected nucleoside with TMS at the 3'-position, **124 β**, bis-Fmoc protected nucleoside with TMS at the 3'-position **125 β**, mono-Fmoc protected nucleoside with no TMS at the 3'-position, **127 β** and bis-Fmoc protected nucleoside with no TMS at the 3'-position, **128 β**. The overall yield for this reaction was *ca.* 82 %. It was thought that during purification by column chromatography, the TMS was partially cleaved. Cleavage of the TMS group in **125 β** was successful using an aqueous solution of potassium fluoride, affording **128 β** in *ca.* 48.6 % yield. This average yield was accounted for by the formation of **127 β** in *ca.* 27 % yield, hence overall yield of *ca.* 76 %. Cleavage of the TMS group is relatively fast, if the reaction mixture is left for too long in aqueous potassium fluoride, the Fmoc groups start to cleave, and hence it was crucial to not prolong reaction duration. **128 β** was then subjected to conventional phosphitylation conditions of 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite and

N,N-diisopropylethylamine (Hünig's base) in DCM affording **129 β** in *ca.* 34 % yield²⁷⁵ (*scheme 2.45*).



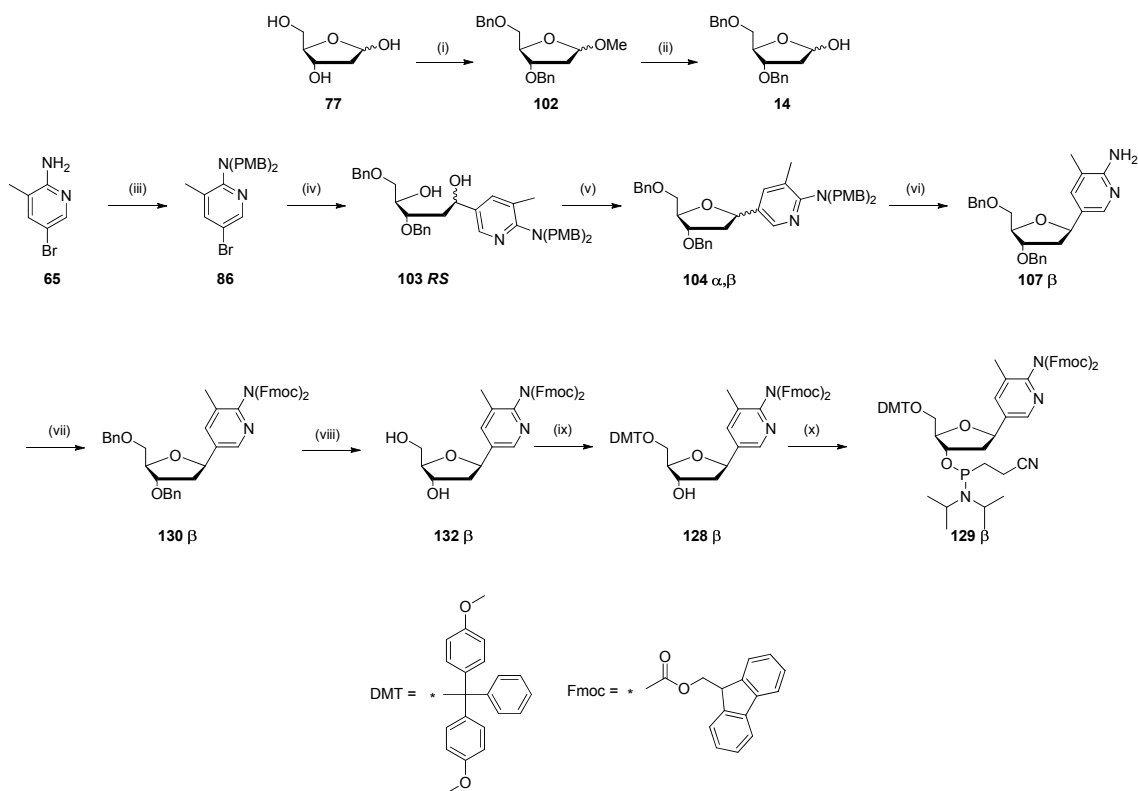
Scheme. 2.45: The optimised synthetic route to ^{Me}P monomer, **129 β**. *Reagents and conditions:* (i) 0.06 eq CH₃COCl, MeOH, r.t, 1 hr, then 4 eq Benzyl Bromide, 4 eq NaH, DMF, 14 hrs, 95 %; (ii) AcOH (80 % in water), reflux, 120 °C, 2 hrs, 60 %; (iii) 2.5 eq 4-methoxybenzyl chloride, 2.5 eq NaH, DMF, r.t, 4 ½ hrs, 76.4 %. (iv) 1.0 eq *n*BuLi, THF, - 35 °C, 30 mins, 0.4 eq **14** in THF, - 30 °C, 1 hr then r.t, 4 ½ hrs, 92 %; (v) 1.6 eq Bu₃P, 1.6 eq TMAD, THF, 0 °C→r.t, 6 ¼ hrs, 88.2 %; (vi) Trifluoroacetic acid: DCM 10:7, r.t, 5 ½ hrs, 98.9 %; (vii) 1.5 eq trifluoroacetic anhydride, pyridine, 0 °C, r.t, 2 ½ hrs, 70 %; (viii) 3 eq BCl₃, DCM, -78 °C, 4 ½ hrs, MeOH, 4 °C, 15 ½ hrs, 98.5 %; (ix) 1.3 eq DMTCl, pyridine, r.t, 2 hrs, 39.5 %; (x) 15 eq TMS-Cl, pyridine, 4 ½ hrs, 4 eq Fmoc-Cl, pyridine, 1 hr 20 mins, water, 1 hr, 82.1 %; (xi) 2 eq KF (1.0M), pyridine, r.t, 1 ½ hrs, 76 %; (xii) 2 eq DIPEA, 1.3 eq 2-cyanoethyl-*N,N*-diisopropylchloro phosphoramidite, DCM, r.t, 2 ½ hrs, 34 %.^{158,202,220,251,253,261,262,275,277}

To summarise, the synthetic route developed here (*scheme. 2.45*) is now a 12-step scheme towards ^{Me}P monomer, **129 β** with overall yield of *ca.* 2.8 % starting from the coupling reaction (*step (iv)*). The free nucleoside, ^{Me}P, **1 β** can be obtained in 8 steps with overall yield of *ca.* 33.3 % starting from the coupling reaction. The overall yield to obtain the perbenzylated lactol, **14** was *ca.* 57 % and that to obtain the fully protected base, **86** was *ca.* 76.4 %. The steps that pulled the overall yield to **129 β** down were the

low yielding tritylation and phosphitylation (*steps (ix) and (xii) respectively*); these can be improved significantly.

It had come to attention that the Fmoc group is an electron-withdrawing group itself and that it may be used in place of TFA in *step (vii)* during the debenzylation, and be able to withstand the Lewis acid conditions.

Successful protection of the exocyclic amino moiety in **107 β** with Fmoc-Cl, in pyridine and acetonitrile afforded the bis-Fmoc protected nucleoside, **130 β** and the mono-Fmoc protected nucleoside, **131 β** respectively, in overall yield of *ca.* 96 %. Debenzylation²⁷⁷ of **130 β** with boron trichloride in DCM at - 78 °C afforded **132 β** (with complete removal of the benzyl groups from both positions of the sugar ring) and **133 β** (3'-*O*-benzyl group present) in an overall yield of *ca.* 79 %. **133 β** could then be put on for further debenzylation reaction to obtain more product, **132 β**. Tritylation¹⁵⁸ at the 5'-position of **132 β** with DMT-Cl in pyridine afforded **128 β**, however the use of triethylamine to neutralise the reaction mixture caused some cleavage of the Fmoc groups. Hence **123 β** was also isolated; the overall yield for this reaction was *ca.* 87 %. This cleavage of the Fmoc groups can be avoided by eliminating the use of triethylamine; the reaction takes place in pyridine and this should be basic enough to mop up the hydrochloric acid produce during the phsphitylation. **123 β** had to then undergo the TMS and Fmoc protection conditions described in *scheme. 2.41*. **128 β** was then subjected to conventional phosphitylation conditions²⁷⁵ of 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite and *N,N*-diisopropylethylamine (Hünig's base) in DCM affording **129 β** in *ca.* 50 % yield (*scheme 2.46*).



Scheme. 2.46: The further optimised synthetic route to ^{Me}P monomer, **129 β**.²⁸² *Reagents and conditions:* (i) 0.06 eq CH₃COCl, MeOH, r.t, 1 hr, then 4 eq Benzyl Bromide, 4 eq NaH, DMF, 14 hrs, 95 %; (ii) AcOH (80 % in water), reflux, 110 °C, 2 hrs, 60 %; (iii) 2.5 eq 4-methoxybenzyl chloride, 2.5 eq NaH, DMF, r.t, 4 ½ hrs, 76.4 %. (iv) 1.0 eq *n*BuLi, THF, - 35 °C, 30 mins, 0.4 eq **14** in THF, - 30 °C, 1 hr then r.t, 4 ½ hrs, 92 %; (v) 1.6 eq Bu₃P, 1.6 eq TMAD, THF, 0 °C→r.t, 6 ¼ hrs, 88.2 %; (vi) Trifluoroacetic acid: DCM 10:7, r.t, 5 ½ hrs, 98.9 %; (vii) pyridine, 8 eq Fmoc-Cl, acetonitrile, 0 °C, r.t, 2 hrs, 95.9 %; (viii) 3 eq BCl₃, DCM, -78 °C, 6 hrs, MeOH, 4 °C, 17 ½ hrs, 78.6 %; (ix) 1.3 eq DMTCl, pyridine, r.t, 3 ½ hrs, 86.8 %; (x) 2 eq DIPEA, 1.21 eq 2-cyanoethyl-*N,N*-diisopropylchloro phosphoramidite, DCM, r.t, 2 hrs, 49.5 %.^{158,202,220,251,253,261,262,275,277}

To summarise, this further optimised synthetic route (*scheme. 2.46*) is a 10-step scheme towards ^{Me}P monomer, **129 β** with overall yield of *ca.* 15.6 % starting from the coupling reaction (*step (iv)*). The phosphitylation step has much room for improvement, and the work-up involved for the DMT protection does not need additional base for neutralising the reaction mixture. Introducing the Fmoc protection earlier in the synthetic route eliminates both the need to protect the exocyclic amino moiety with TFA protection, and the need to protect the 3'-position of the sugar ring with TMS.

The ^{Me}P monomer, **129 β** was incorporated into various oligonucleotides as protonated Cytidine equivalents and compared with other Cytidine-analogues (5-methylcytosine

(^{Me}C), 2'-*O*-methoxy-3-methyl-2-aminopyridine (^{Me}P_{OMe}) and 2'-*O*-aminoethyl-5-methylcytosine (^{Me}C_{AE}). Successful formation of the putative triplexes was confirmed *in vitro* using biophysical techniques; UV-melting, photo cross linking and footprinting studies, which is discussed further in Chapter 3.

To conclude, a concise, efficient route (*scheme. 2.46*), which has been extensively optimised to provide a successfully protected 3-methyl-2-aminopyridine (^{Me}P) monomer, **129 β** has been developed. To arrive at this efficient route many different approaches had to be explored and some eliminated.

In the future, this synthetic route can be further optimised by attempting to improve β-glycosylation *via* the suggestions mentioned above.

CHAPTER 3

Results and Discussion

Biophysical and Biological Studies

3. BIOPHYSICAL AND BIOLOGICAL STUDIES

3.1 INITIAL STUDIES ON ^{Me}P

3.1.1 *IN VITRO* STUDIES

3.1.1.1 FOOTPRINTING STUDIES

THE EFFECT OF ^{Me}P ON THE BINDING AFFINITY OF AN OLIGONUCLEOTIDE TO THE GFP CODING REGION IN *C. ELEGANS* AS A FUNCTION OF PH

DNase I footprinting was initially carried out to examine whether 3-methyl-2-amino-5-(2'-deoxy-β-D-ribofuranosyl)pyridine (^{Me}P), within the oligo **OL1**, can achieve high affinity recognition of a target site (*fig. 3.1*), at pH values up to physiological pH. The target site was a 14-base oligopurine tract interrupted by a T; it is part of the *gfp* coding gene, and was obtained from a pUC19 clone.

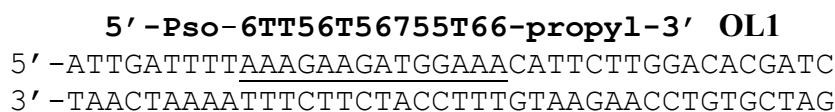


Fig. 3.1: Sequence of the 14 base-pair target site (underlined) within part of the *gfp* coding gene, which was obtained from a pUC19 clone. TFO **OL1** shown in bold; where **5** = ^{Me}P; **6** = T_{AE}; **7** = S_{ME} and **Pso** = Psoralen (see page 280-281 for the respective structures of these base analogues).

Psoralen was also incorporated into the TFO. If footprinting of this oligo proved successful at physiological pH, the psoralen cross linking function could then be utilized to enhance triplex binding even further.

A restriction fragment labelled at the 3'-end with ³²P, containing the target site (outlined in *fig. 3.1*) was incubated overnight at room temperature with increasing concentrations of the triple helix forming oligonucleotide **OL1** (ranging from 0.03 μM to 10 μM). DNase I generated reaction products were separated by denaturing gel electrophoresis. After fixing, drying, and storing the gel overnight in a phosphorimager screen, it was scanned for visualisation of the footprint.

Figs. 3.2 (A)-(C) show representative footprinting patterns at part of the *gfp* coding gene target site, obtained at pH 5.5, 6.2 and 7.0, respectively. It should be noted, that when cloning of the target sequence was carried out, the resultant clone contained a dimer of this sequence with the two copies oriented in opposite directions, enabling simultaneous visualisation of the purine (upper site) and pyrimidine (lower site) strands. There exists a third region which appears to be another footprint; it is in fact part of the second footprint, which is split into two.

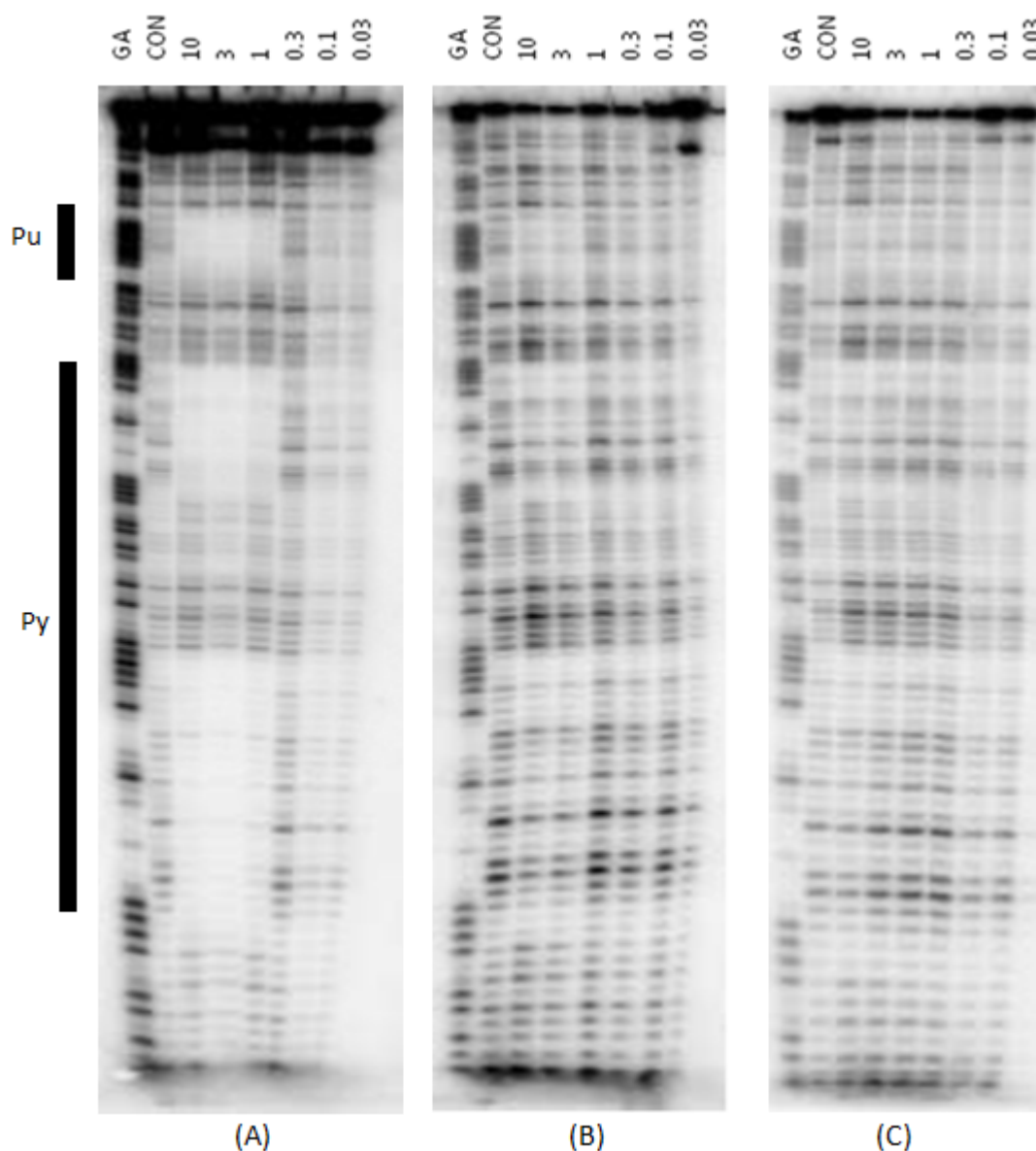


Fig. 3.2: DNase I footprints for the interaction of TFO **OL1** with the 14 base-pair target site within part of the *gfp* coding gene. Oligonucleotide concentrations (μM) are indicated at the top of each gel lane. The reactions were performed in: (A) 50mM sodium acetate, pH 5.5; (B) 10 mM PIPES, pH 6.2, containing 50mM NaCl; and (C) 10 mM Tris-HCl, pH 7.0, containing 50 mM NaCl. Tracks labelled “GA” are sequencing lanes that are specific for purines (G + A); tracks labelled “CON” represent control where there was no TFO present. The solid bars indicate the location of the target site; ‘Pu’ indicates the purine-rich strand of the target site; ‘Py’ indicates the pyrimidine-rich strand.

At pH 5.5 it can be clearly seen that **OL1** prevents DNase I activity down to a concentration of 1 μM (fig. 3.2). At pH 6.2 there exists some attenuation of the cleavage products for **OL1** concentrations of 10 μM and 3 μM , suggesting that the TFO also reduces DNase I activity here as well. However, for the lower concentrations of **OL1** at

pH 6.2, and for all concentrations at pH 7.0, no footprint is evident, which was later accounted for (see below).

Analysis by negative mode electrospray mass spectrometry showed that oligonucleotide **OL1** contained the peak for the approximate expected mass of the oligo of 5199 (M-H)⁻. However, peaks at 5228, 5256, 5284 etc, were also present, and these corresponded to (M-H + n(28))⁻ where n = 1,2,3 & 4. It was thought that this extra mass of 28 was originating from an extra carbonyl group being present. This was likely to have been derived from the *N,N*-dimethylformamidine protecting group used to protect the exocyclic amino function at the 2- position in ^{Me}**P**. It had been predicted that the dimethylformamidine group had undergone incomplete deprotection (hydrolysis) under standard oligonucleotide deprotection conditions (concentrated aqueous ammonia, room temperature, 24 hours). The formamide derivative is the intermediate formed during deprotection of the *N,N*-dimethylformamidine²⁸³ (*fig. 3.3*). For complete cleavage, harsher conditions were required. If temperatures had been elevated during the deprotection of **OL1**, the *N,N*-dimethylformamidine protecting group may have fully cleaved. However, heating the ammonia solution to high temperatures such as 55 °C would have caused degradation of the Psoralen that had been incorporated into the oligo.²⁸⁴ Hence, these harsher conditions could not be employed during oligonucleotide deprotection. Therefore, the use of the formamidine group as protection for the exocyclic amino moiety was discontinued.

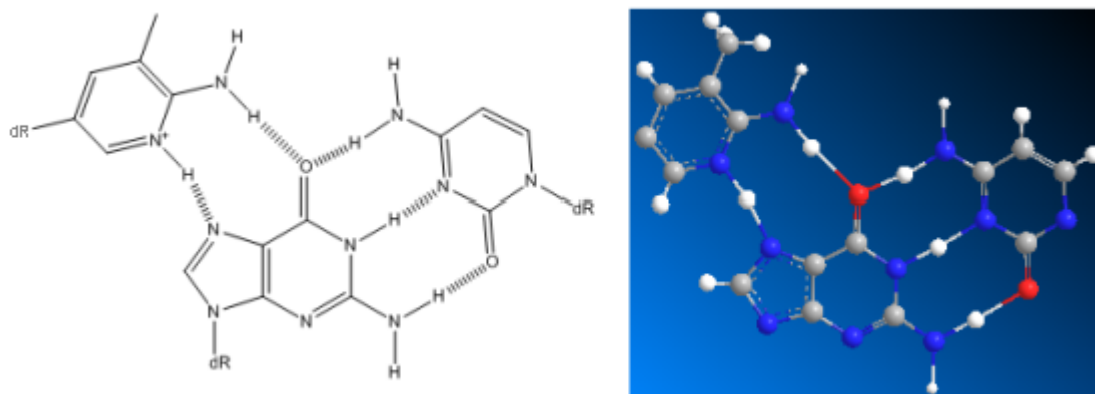


Fig. 3.4: $^{\text{Me}}\text{P}\cdot\text{GC}$ triplet shown in 2-D view (left) and as a ball-and-stick 3-D view (right); where red represents oxygen, blue represents nitrogen, grey represents carbon, and white represents hydrogen.

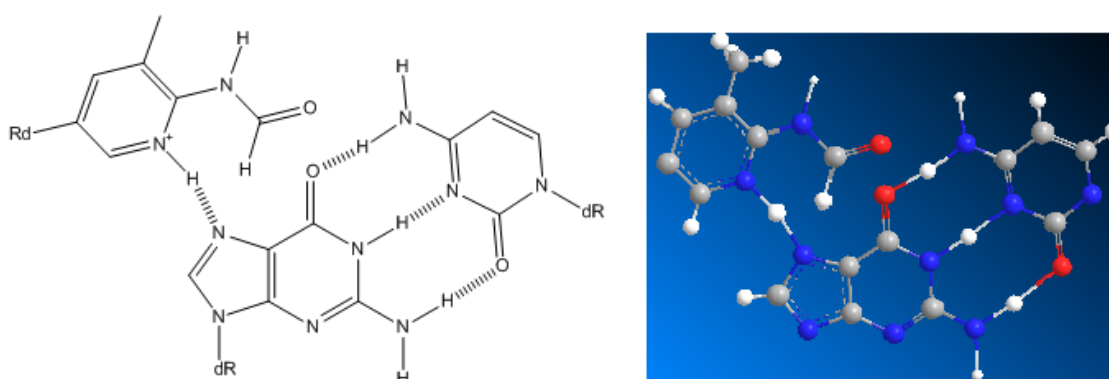


Fig. 3.5: The disrupted $^{\text{Me}}\text{P}\cdot\text{GC}$ triplet that cannot form the appropriate Hoogsteen hydrogen bonds required for triplex formation, shown in 2-D view (left) and as a ball-and-stick 3-D view (right); where red represents oxygen, blue represents nitrogen, grey represents carbon, and white represents hydrogen.

3.2 COMPARISON OF CYTOSINE-NUCLEOSIDE-ANALOGUES IN TFOS

3.2.1 *IN VITRO* STUDIES

3.2.1.1 UV-MELTING – THERMAL DENATURATION STUDIES

The Cytosine nucleoside analogues (C-analogue); 3-methyl-2-amino-5-(2'-deoxy- β -D-ribofuranosyl)pyridine ($^{\text{Me}}\text{P}$), 3-methyl-2-amino-5-(2'-O-methyl- β -D-ribofuanosyl)pyridine ($^{\text{Me}}\text{P}_{\text{OMe}}$), and 5-methyl-(2'-O-aminoethyl- β -D-ribofuranosyl)cytosine ($^{\text{Me}}\text{C}_{\text{AE}}$) were incorporated into TFOs **OL3**, **OL2** and **OL5** respectively (The analogues $^{\text{Me}}\text{P}_{\text{OMe}}$ and $^{\text{Me}}\text{C}_{\text{AE}}$ were synthesised by Mr. C. Lou and

initial triplex to duplex plus TFO transition, followed by duplex dissociation. The melting curve which was obtained for **OL3**/^{Me}**P** and **OL4**/^{Me}**C** at pH 6.2, revealed a single transition, with T_m value of the triplex in the same range as that of the underlying duplex, demonstrating that these triplexes were highly stable at this pH.

COMPARISON OF THE C-NUCLEOSIDE ANALOGUES AND 5-METHYL CYTOSINE

	pH 6.2	pH 6.6	pH 7.0	pH 7.5	pH 8.0
OL2 / ^{Me} P _{OMe}	60.06/60.27	53.79/56.76	45.67/46.08	38.54/38.95	30.20/29.76
OL3 / ^{Me} P	>70.0/>70.0	69.43/69.73	60.69/60.53	52.47/53.13	44.75/44.90
OL4 / ^{Me} C	69.77/69.94	61.71/61.50	53.73/54.05	44.08/44.41	35.04/32.10
OL5 / ^{Me} C _{AE}	64.77/63.43	54.78/55.89	47.24/47.05	36.07/36.39	25.03/25.66

Table. 3.1: UV triplex melting experiments for the C-analogues in 10 mM sodium phosphate, 200 mM NaCl, pH 6.2-8.0, using the standard melt program (20-80-20 °C at 0.5 °C/min). Temperature values for annealing (T_a) and those for melting (T_m) are given in that order respectively, and quoted in °C. Temperature values were determined from a single UV-melting cycle.

	pH 6.2	pH 6.6	pH 7.0	pH 7.5	pH 8.0
OL2 / ^{Me} P _{OMe}	0.21	2.97	0.41	0.41	0.44
OL3 / ^{Me} P	0.00	0.30	0.16	0.66	0.25
OL4 / ^{Me} C	0.17	0.21	0.32	0.33	2.94
OL5 / ^{Me} C _{AE}	1.34	1.11	0.19	0.32	0.63

Table. 3.2: Hysteresis (the difference between the melting (T_m) and annealing temperatures (T_a)) observed, in °C.

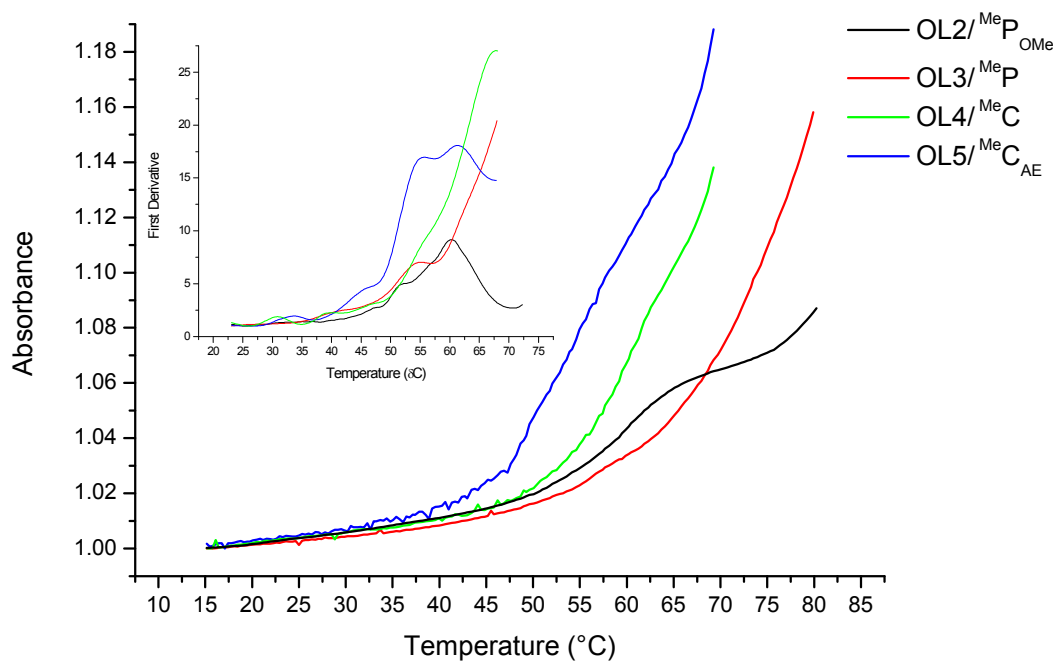


Fig. 3.7: UV-melting curves and the first derivatives for the triplexes formed between **OL9** and: **OL2/^{Me}P_{OMe}**, **OL3/^{Me}P**, **OL4/^{Me}C** and **OL5/^{Me}C_{AE}** at pH 6.2.

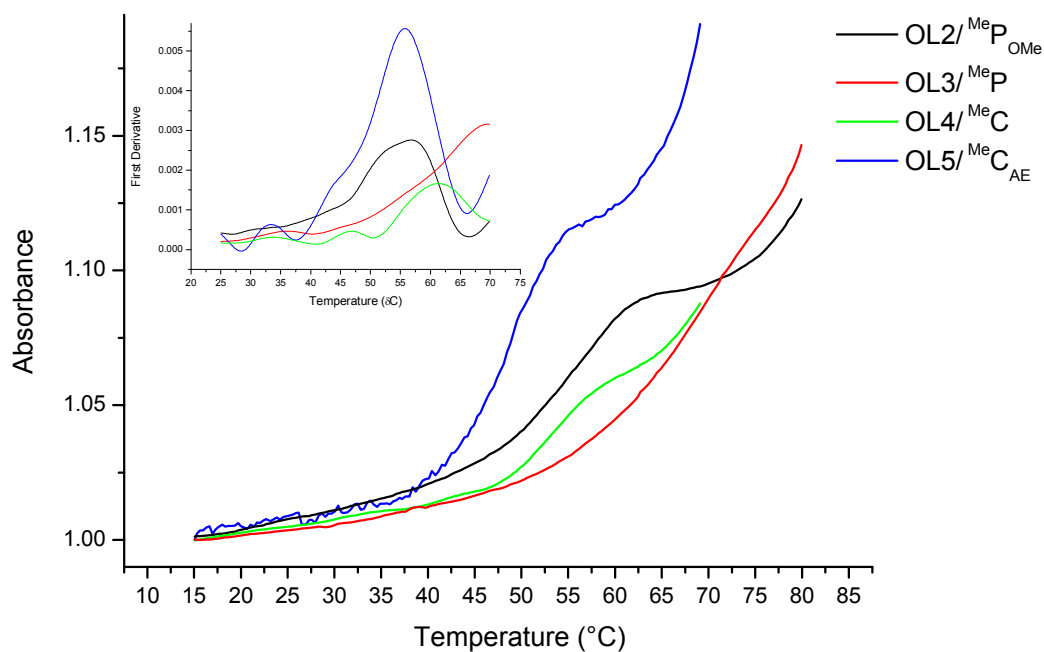


Fig. 3.8: UV-melting curves and the first derivatives for the triplexes formed between **OL9** and: **OL2/^{Me}P_{OMe}**, **OL3/^{Me}P**, **OL4/^{Me}C** and **OL5/^{Me}C_{AE}** at pH 6.6.

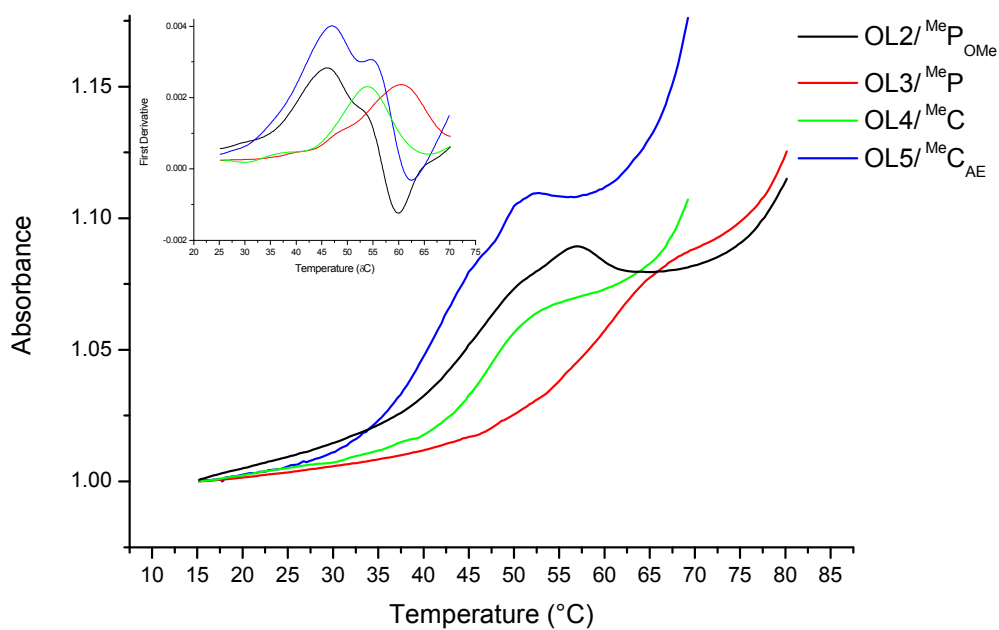


Fig. 3.9: UV-melting curves and the first derivatives for the triplexes formed between **OL9** and: **OL2/^{Me}P_{OMe}**, **OL3/^{Me}P**, **OL4/^{Me}C** and **OL5/^{Me}C_{AE}** at pH 7.0.

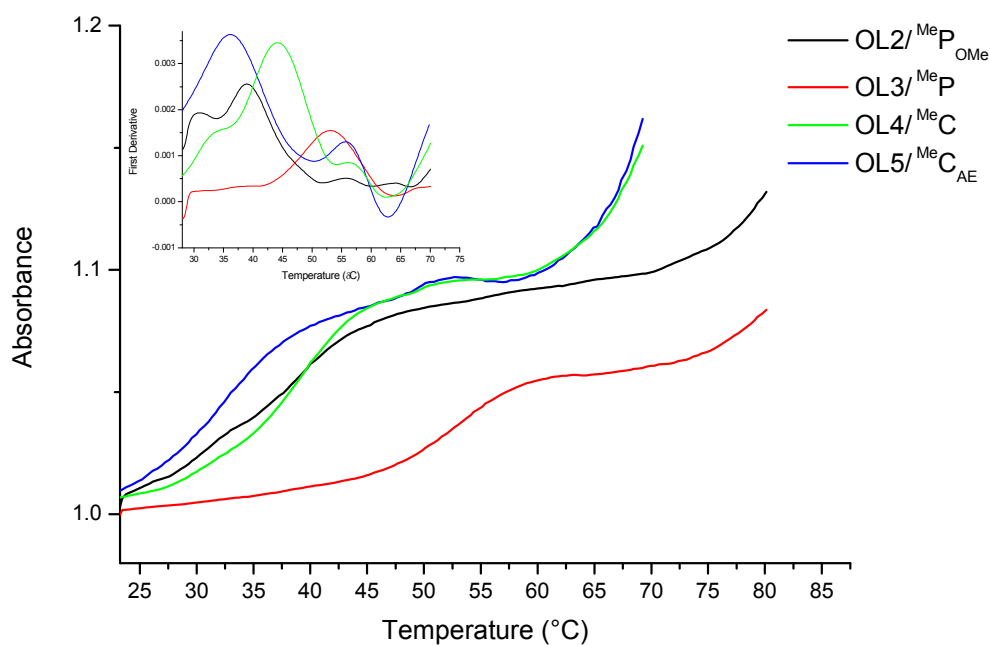


Fig. 3.10: UV-melting curves and the first derivatives for the triplexes formed between **OL9** and: **OL2/^{Me}P_{OMe}**, **OL3/^{Me}P**, **OL4/^{Me}C** and **OL5/^{Me}C_{AE}** at pH 7.5.

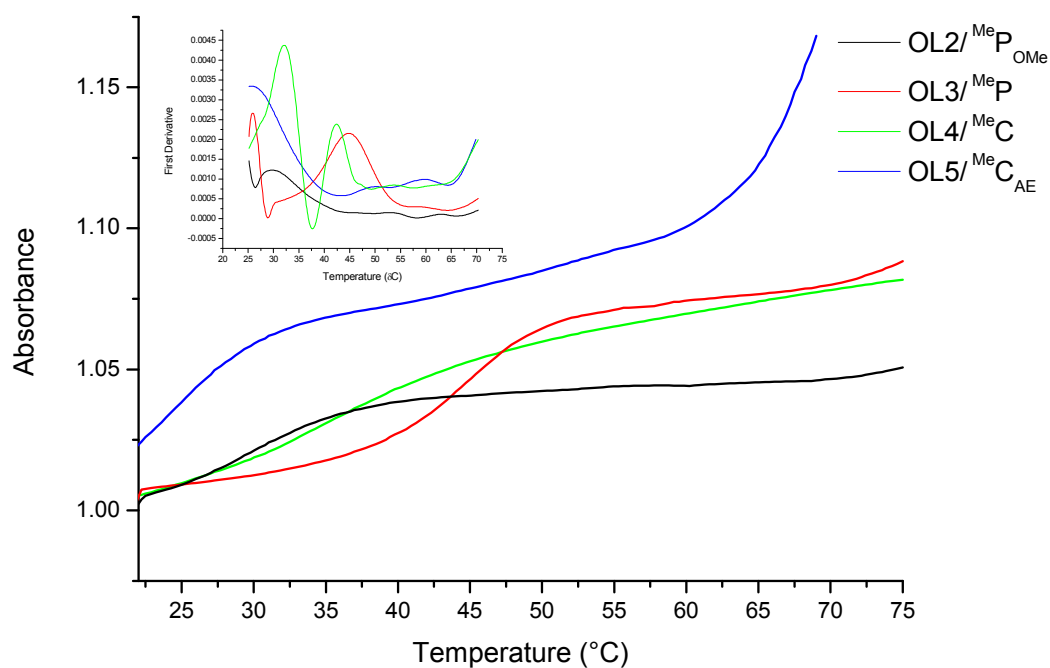


Fig. 3.11: UV-melting curves and the first derivatives for the triplexes formed between **OL9** and: **OL2/MeP_{OMe}**, **OL3/MeP**, **OL4/MeC** and **OL5/MeC_{AE}** at pH 8.0.

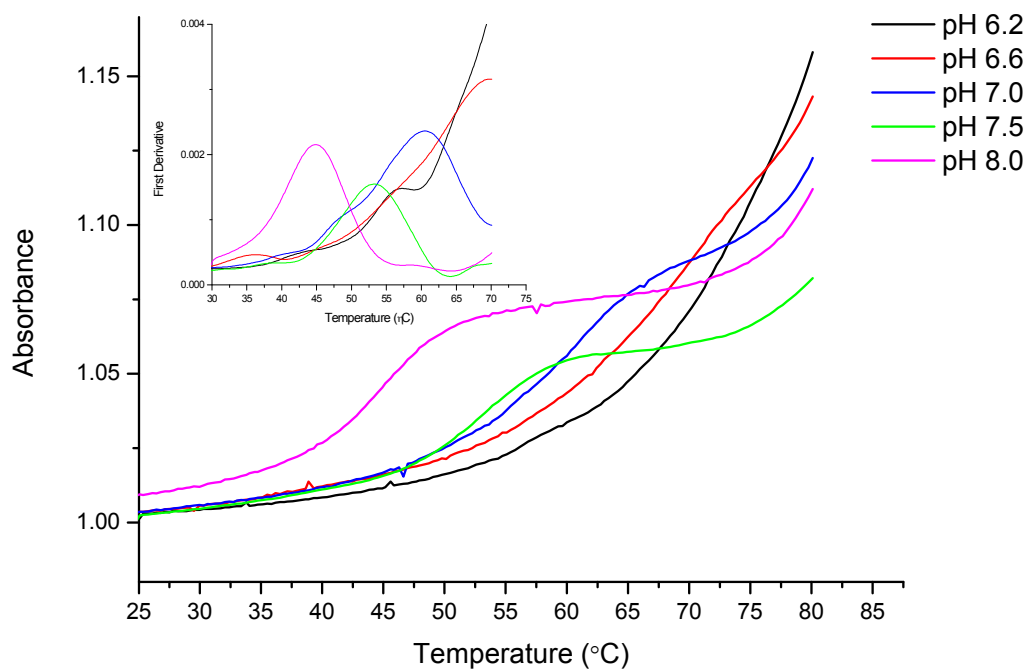


Fig. 3.12: UV-melting curves and the first derivatives for the triplex formed between **OL3/MeP** and **OL9** hairpin duplex at pH 6.2-8.0.

From the data in *table. 3.1*, it can be observed that in many cases, the thermal dissociation temperatures (T_m) are higher than the association/annealing (T_a) temperatures. From the data presented in *table. 3.2*, in general, hysteresis is dependent on pH within the range pH 6.2-8.0. Overall, as the pH increases, so does the hysteresis (there exists a few results which do not follow the trend). For all the oligos studied, at pH 6.2, with the exception of **OL5**/^{Me}C_{AE}, the degree of hysteresis encountered is at its lowest. The heating and cooling curves for **OL3**/^{Me}P are more or less superimposed at pH 6.2.

From the data in *table. 3.1*, it appears that at all pH, **OL3** in which all ^{Me}C were replaced by the C-nucleoside ^{Me}P, shows an increase in T_m relative to **OL4**/^{Me}C, containing ^{Me}C. The increase in stability is most pronounced above pH 7.0. The average increase in T_m /modification for ^{Me}P is 2.18 °C at pH 7.5, and 3.2 °C at pH 8.0, relative to ^{Me}C. The results clearly demonstrate that, in the pH range 6.2-8.0, within the TFO sequence under examination, the presence of ^{Me}P had a positive effect on triplex stability. From the oligos studied here, at all pH, **OL3**/^{Me}P gives the highest triplex melting temperatures. Hence, by replacing ^{Me}C in a TFO with ^{Me}P, which is ~ 2 pK_a units more basic than ^{Me}C, results in an enhanced affinity to a double stranded DNA target, over the pH range 6.2-8.0.

The results also conclude that the presence of either a 2'-O-methyl or 2'-O-aminoethyl modification is ineffective in enhancing triplex stability, they are actually destabilising, as reported previously by Hildbrand *et al.*¹⁵⁸ (*discussed further below*). The decrease in stability for **OL2**, containing ^{Me}P_{OMe}, when compared to **OL4** containing ^{Me}C, is most pronounced at pH 6.2; the average decrease in T_m /modification is 2.4 °C. Whereas, the decrease in stability for **OL5**, containing ^{Me}C_{AE}, when compared to **OL4**/^{Me}C, is most pronounced above pH 7.0; the average decrease in T_m /modification is 2.01 °C at pH 7.5.

This is contrary to the results reported by R. W. Roberts and D. M. Crothers.²⁸⁵ They state that RNA third strands have been recognised to form more stable pyrimidine triplexes than the corresponding DNA third strands; this observed variability must lie in the chemical differences between RNA and DNA: the 2'-OH.²⁸⁶ Both **OL2**/^{Me}P_{OMe} and **OL5**/^{Me}C_{AE} contain Cytosine analogues in which their sugar resembles an RNA-like sugar conformation.

For all pH buffers 6.2-8.0, the general trend observed in terms of thermal triplex stability was as follows:

$$^{\text{Me}}\text{P} > ^{\text{Me}}\text{C} > ^{\text{Me}}\text{P}_{\text{OMe}} > ^{\text{Me}}\text{C}_{\text{AE}}$$

It was expected that the thermal stability of **OL3**/^{Me}P would be greater than that of **OL4**/^{Me}C, which is displayed in the results. However, it was also expected that the trend would be close to that below:

$$^{\text{Me}}\text{C}_{\text{AE}} > ^{\text{Me}}\text{P}_{\text{OMe}} > ^{\text{Me}}\text{P} > ^{\text{Me}}\text{C}$$

A DNA single strand alone usually adopts a C^{2'}-*endo* conformation. NMR studies^{287,288} and polynucleotide fibre diffraction analyses²⁸⁹ of DNA triplexes show that the sugars of the third strand have RNA-like, C^{3'}-*endo* puckering.⁶⁵ The 2'-*O*-methyl-3-methyl-2-aminopyridine monomer, ^{Me}P_{OMe} was designed to improve the effectiveness and strength of binding of the TFO in which it was incorporated, in comparison to the deoxy version. This 2'-OMe ribose modification has been reported²⁸⁶ to impart a C^{3'}-*endo* conformation, which is optimal for triplex formation by pyrimidine oligonucleotides, imposing a lower degree of distortion on the underlying duplex than third strands containing deoxyribose based nucleosides. This has the salutary effect of reducing some of the entropic barriers to triplex formation.²⁸⁶ The methoxy groups enhance the rigidity of the triple stranded structure,⁶⁵ and van der Waals interactions exist between the methyl groups and the i-1 purine residues in the duplex.²⁸⁶

The results obtained in this study agree to some extent with that reported by Hildbrand *et al.*¹⁵⁸ In comparison to the deoxy nucleoside, ^{Me}P, the ribose nucleoside, ^{Me}P_{OMe} does not enhance triplex stability. The TFO sequence studied by Hildbrand *et al.* contained five ^{Me}P_{OMe}, and for conditions above pH 6.5, no affinity to the target duplex was observed. For the TFO sequence containing **OL2**/^{Me}P_{OMe}, affinity is observed above pH 6.5, however it is weak in comparison to the affinity of the TFOs that contain either ^{Me}P or ^{Me}C. The difference between the two studies; in this case the TFO under investigation contains a uniform ribose backbone, whereas the TFOs studied by Hildbrand *et al.*¹⁵⁸ had a heterogeneous backbone composition (composed of both 2'-deoxy and 2'-*O*-methyl ribose units). As suggested by Hildbrand *et al.*¹⁵⁸ the

heterogeneous oligonucleotides behave differently from their pure ribo and deoxyribo sequences, and this could be due to differences in preferred conformation of the third strand or the triplex itself.¹⁵⁸ In both this study and that carried out by Hildbrand *et al.*,¹⁵⁸ the replacement of **MeP** with **MeP_{OMe}** results in a distinct loss in triplex stability, although there is no relevant difference in the pK_a of these two nucleosides.

The 2'-*O*-aminoethyl modified base analogues are designed for dual recognition. These analogues are able to recognise duplex DNA through base–base contacts, as well as by the formation of a salt bridge between the protonated amino function of the oligonucleotide and an oxygen of the negatively charged phosphodiester unit of the target purine strand (*fig. 3.14*).²⁹⁰ The presence of the 2'-*O*-aminoethyl side chains also contributes actively to the nucleation-zipping association process. Specific contacts between the 2'-*O*-aminoethyl and the phosphate backbone of the DNA duplex are able to form, and this provides for a dramatic increase in the binding affinity as well as in the association rate constant.²⁹¹ The sugar conformation is *C3'-endo-2'-exo*, and the underlying duplex encounters only slight distortion.²⁸⁴

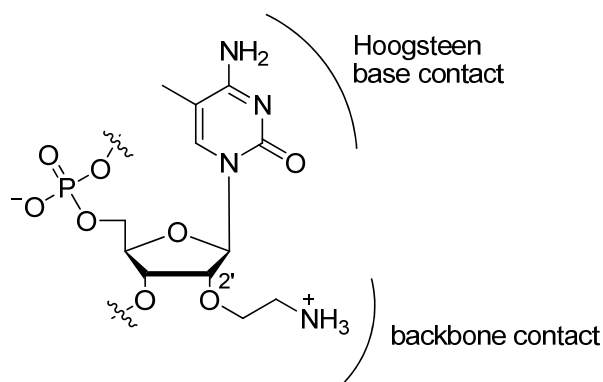


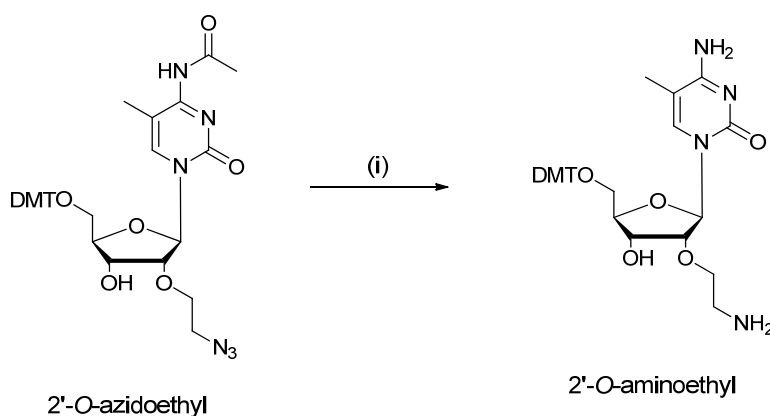
Fig. 3.14: The structure of the **MeCAE**, showing the region where Hoogsteen hydrogen bonds form and where interaction with the phosphodiester backbone occurs with the purine strand of the underlying duplex.²⁹¹

Due to the combination of its positive charge and RNA-like sugar conformation, **MeCAE** was expected to have the most profound effect on the kinetics of triplex formation for the TFO in which it was incorporated. It was also expected that the resultant complex would have a greater stability at physiological pH and low Mg^{2+} concentration, than any

of the other three Cytosine analogues. Hence, it was a surprise when the results obtained did not reflect this.

It was later determined that the 2'-*O*-aminoethyl group was being acetylated during the synthesis of ^{Me}C_{AE}. This was evident by the presence of the peak in the oligonucleotide MS at 5851 (*see page 278*), corresponding to the mass of the oligo plus four additional acetyl groups (+172); one per ^{Me}C_{AE}. A peak at 5678 was that expected in the absence of acetyl migration.

It was determined from 2-D NMR studies that the acetyl protecting group was migrating from the exocyclic amino function during the Staudinger reaction,²⁹² in which the 2'-*O*-azidoethyl group was reduced to 2'-*O*-aminoethyl (*scheme. 3.15*). The COSY NMR displays correlation between the protons of one of the CH₂ in the 2'-*O*-aminoethyl group, with the protons in the CH₃ of the acetyl group.



Scheme. 3.15: The Staudinger reaction. *Reagents and conditions:* (i) 2 eq Triphenyl phosphine, 5 eq Water, THF, 45 °C, 14 hrs.²⁹²

It was thought that for the acetyl migration to occur, the exocyclic amino group and the 2'-*O*-aminoethyl chain would have to be in close proximity in space. Once the azide functional group had been reduced to amine functionality to give a primary alkyl amine, this region of the molecule was highly nucleophilic and therefore more reactive, in comparison to the exocyclic amine attached at the C-4 position of the aromatic ring in the 5-methylcytosine base. The trigonal pyramidal arrangement of bonds around nitrogen is shallower in aryl amines vs. alkyl amines. This is a result of resonance

delocalisation of the lone pair into the aromatic π system; it is this delocalisation that is responsible for the decreased nucleophilicity of aryl vs. alkyl amines. The acetyl migration is thought to have occurred rapidly after reduction of the azide (*fig. 3.16*).

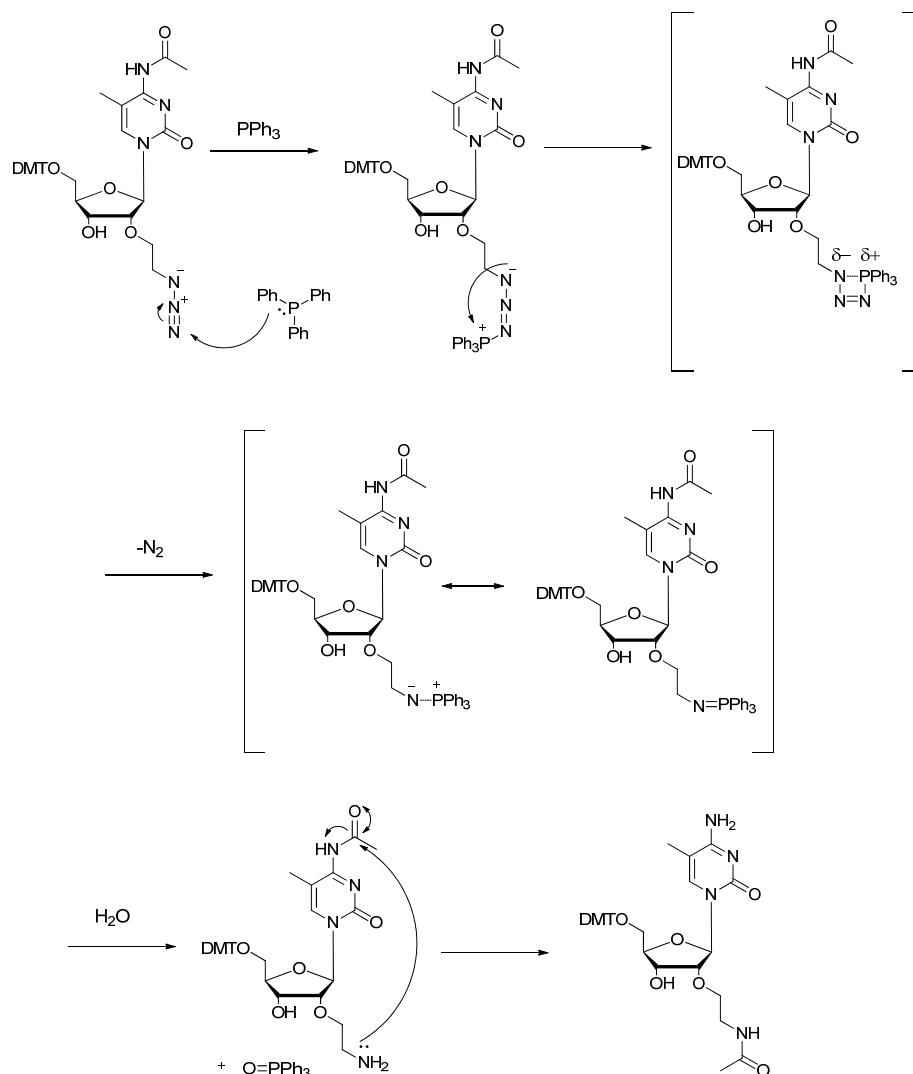


Fig. 3.16: The mechanism of the Staudinger reaction and the rapid acetyl migration that follows.

Once the 2'-*O*-aminoethyl group had been acetylated, it was inert to aqueous ammonia which was used in oligonucleotide deprotection. The ammonia required to deprotect the acetyl group from the exocyclic amino group was not nucleophilic enough to cleave the acetyl from the 2'-*O*-aminoethyl, harsher conditions were required.

Due to this problem being unforeseen at the start of the synthesis, the synthetic development of the $^{\text{Me}}\text{C}_{\text{AE}}$ monomer was monitored each step by 1D NMR and MS. The

molecular analytical results were misleading due to the presence of two amino functionalities; it was difficult to differentiate using 1D NMR, which position had been acetylated. On discovering that acetyl migration is possible, more precise monitoring methods need to be put in place to determine exactly the position of the acetyl group (e.g. detailed NMR studies or x-ray crystallography). $^{Me}C_{AE}$ monomer synthesis is currently being modified to prevent this unwanted acetyl migration from occurring.

Pyrimidine motif triplexes are known for their sensitivity to increase in pH. The decline in stability of triplexes in the pyrimidine motif as the pH is increased is ascribed to the loss of protonation at the nitrogen within the base of the C-analogues. Although there is also an accompanying loss of a hydrogen bond, it has been shown that the loss of the positive charge is the more critical factor.²⁸⁶

Analysis of the triplexes formed by **OL2-5** as a function of pH (6.2-8.0) (*fig. 3.17*; see also *fig. 3.12* for the melting curves and first derivatives of **OL3**/ ^{Me}P) showed the predictable decline in T_m and drop in triplex stability as the pH is increased.

Both **OL2**/ $^{Me}P_{OMe}$ and **OL3**/ ^{Me}P showed a more pronounced decline in T_m between pH 6.6-7.0; -10.68 °C for **OL2**/ $^{Me}P_{OMe}$ and -9.20 °C for **OL3**/ ^{Me}P . Whereas, between pH 7.0 and 8.0, the decline in T_m was more gradual as the pH increased. This was illustrated by the shallower T_m differential in this section of the profiles; **OL2**/ $^{Me}P_{OMe}$: -7.13 °C between pH 7.0-7.5 and -9.19 °C between pH 7.5-8.0; **OL3**/ ^{Me}P : -7.4 °C between pH 7.0-7.5 and -8.23 °C between pH 7.5-8.0. Hence, these two oligonucleotides are at their most sensitive within the pH range 6.6-7.0, while above pH 7.0, the sensitivity decreases somewhat. This was expected; the pK_a s of ^{Me}P and its 2'-*O*-methyl derivative, $^{Me}P_{OMe}$ is ~ 6.33 (pK_a range 6.0-7.0). Hence, the pH sensitivity of triplex melting will be most pronounced in this pH range (pH 6.2-7.0). This is because exactly half of the nucleosides are protonated and the other half unprotonated. On moving away from the pK_a range of these bases to higher pH, these bases will be fully deprotonated, and therefore the triplex T_m will be much less sensitive to changes in pH.

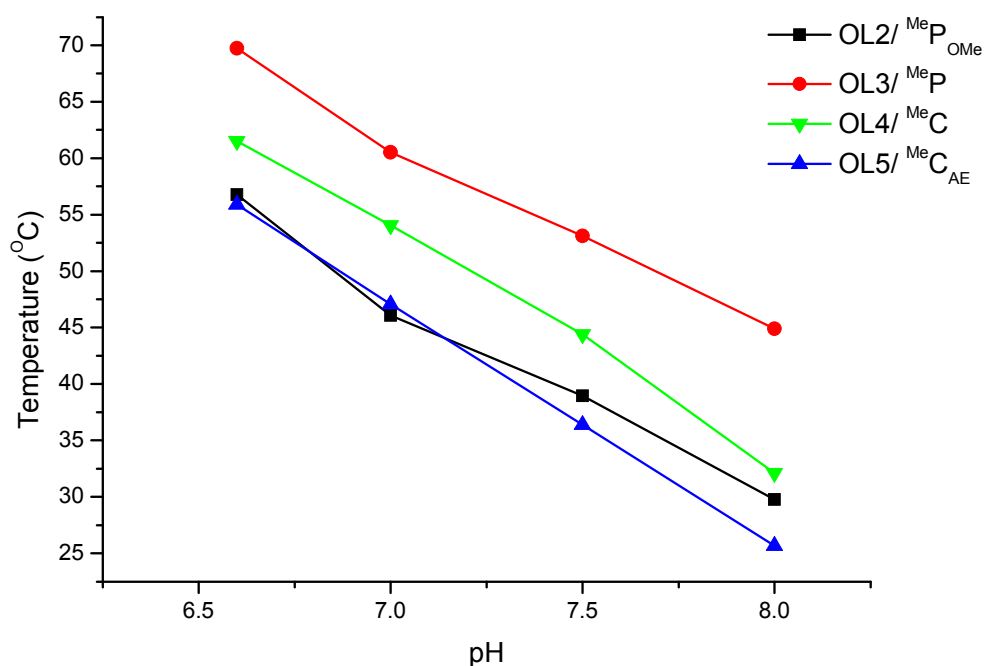


Fig. 3.17: T_m vs. pH profile, representing the drop in T_m as a function of pH for the triplexes formed between **OL9** and **OL2/^{Me}P_{OMe}**, **OL3/^{Me}P**, **OL4/^{Me}C** and **OL5/^{Me}C_{AE}**.

A shallower T_m differential was observed for **OL4/^{Me}C** and **OL5/^{Me}C_{AE}** between pH 6.6-7.0; -7.45 °C for **OL4/^{Me}C**, and -8.84 °C for **OL5/^{Me}C_{AE}**. Over the pH range 7.0-8.0, a slightly larger T_m differential was observed than that for **OL2/^{Me}P_{OMe}** and **OL3/^{Me}P**. For **OL4/^{Me}C**, -9.64 °C T_m differential was observed between pH 7.0-7.5, and -12.31 °C between pH 7.5-8.0. For **OL5/^{Me}C_{AE}**, -10.66 °C T_m differential was observed between pH 7.0-7.5, and -10.73 °C between pH 7.5-8.0. Hence, **OL4/^{Me}C** and **OL5/^{Me}C_{AE}** are at their most sensitive in the pH range 7.0-8.0.

Overall, when the pH was increased from 6.6-8.0, the T_m differential was: -27.00 °C for **OL2/^{Me}P_{OMe}**; -24.83 °C for **OL3/^{Me}P**; -29.40 °C for **OL4/^{Me}C**; and -30.23 °C for **OL5/^{Me}C_{AE}**. Thus, from these results, it can be stated that, an oligo containing ^{Me}P encounters the least sensitivity to increase in pH, whereas an oligo containing ^{Me}C_{AE} encounters the most sensitivity to increase in pH. Hence, the presence of ^{Me}P in a TFO, in comparison to the other three C-analogues, partially ameliorates the destabilising effects of pH in the physiological range. This lower pH dependence of the T_m values for ^{Me}P containing oligo in comparison to ^{Me}C containing oligo is a direct consequence of

the higher pK_a . The observation that $^{Me}C_{AE}$ encounters the most sensitivity to pH is a confirmation that the 2'-*O*-aminoethyl group was acetylated as described above; the pK_a of this amino group is 8.5-9.0. Hence it would still be protonated at physiological pH and higher, and this positive charge would partially compensate for the loss of the charge on the base of the C-analogue, as the pH moves to the physiological range. Therefore, the TFO with $^{Me}C_{AE}$ (**OL5**) should not encounter the most sensitivity to increase in pH.

3.2.1.2 FOOTPRINTING STUDIES

DNase I footprinting was used to make a comparison between the oligos: **OL2**/ $^{Me}P_{OMe}$, **OL3**/ ^{Me}P and **OL4**/ ^{Me}C . This was to observe which C-analogue; $^{Me}P_{OMe}$, ^{Me}P or ^{Me}C respectively, would enable the TFO within which it is incorporated to have the greatest affinity at physiological pH, to the target site, a 15 base-pair oligopurine tract interrupted by a cytosine. The target site is part of the *slo-1 promoter* gene in *C. elegans* (fig. 3.18).

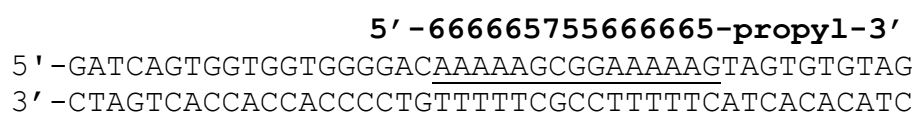


Fig. 3.18: DNase I footprinting reaction; Sequence of the 15 base-pair target site (underlined) part of the *slo-1 promoter* gene. TFO shown in bold; **5** = $^{Me}P/^{Me}P_{OMe}/^{Me}C$ (**OL3/OL2/OL4** respectively) ; **6** = T_{AE} ; and **7** = S_{ME} (see page 280-281 for the respective structures of these base analogues).

A restriction fragment labelled at the 3'-end with ^{32}P , containing the target site outlined in fig. 3.18, was incubated at room temperature separately with the triple helix forming oligonucleotides **OL2**/ $^{Me}P_{OMe}$, **OL3**/ ^{Me}P and **OL4**/ ^{Me}C (increasing concentrations ranging from 0.01/0.03 μM to 10 μM). DNase I generated reaction products were separated by denaturing gel electrophoresis, and after fixing, drying, and storing overnight in a phosphorimager screen, the footprint was visualised by scanning.

Figs. 3.19-3.21 show representative footprinting patterns at a region of the *slo-1 promoter* gene target site, obtained at pH 5.0, 7.0 & 7.5 respectively. As previously stated, it should also be noted here, that when cloning of the target sequence was carried out, the resultant clone contained a dimer of this sequence with the two copies oriented

in opposite directions, enabling simultaneous visualisation of the purine (upper site) and pyrimidine (lower site) strands. There exists a third region which appears to be another footprint; it is in fact part of the second footprint, which is split into two.

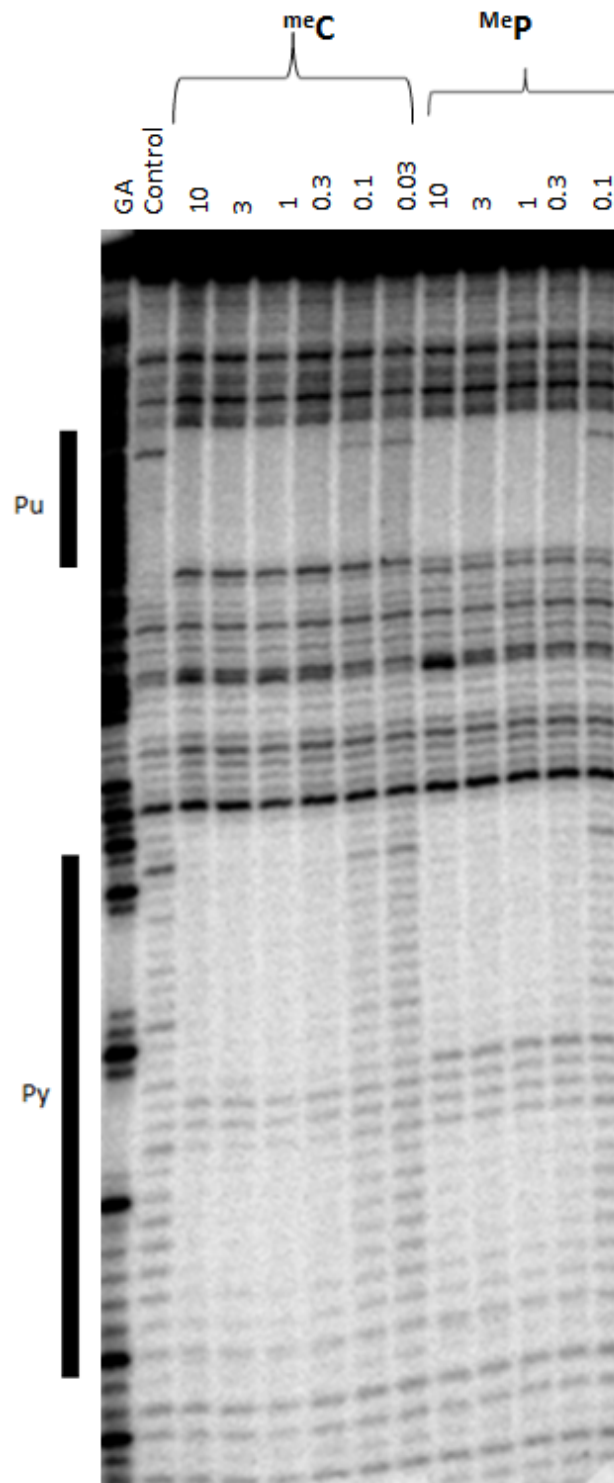


Fig. 3.19: DNase I cleavage patterns produced by **OL3/MeP** (right) and **OL4/MeC** (left) with the 15 base-pair oligopurine target site interrupted by a cytosine, that exists within the *slo-1 promoter* gene region. Oligonucleotide concentrations (μM) are indicated at the top of each gel lane. The reaction was performed in 50mM sodium acetate, pH 5.0. Track labelled “GA” is the sequencing lane that is specific for purines (G + A). The solid bars indicate the location of the target site; ‘Pu’ indicates the purine-rich strand of the target site; ‘Py’ indicates the pyrimidine-rich strand.

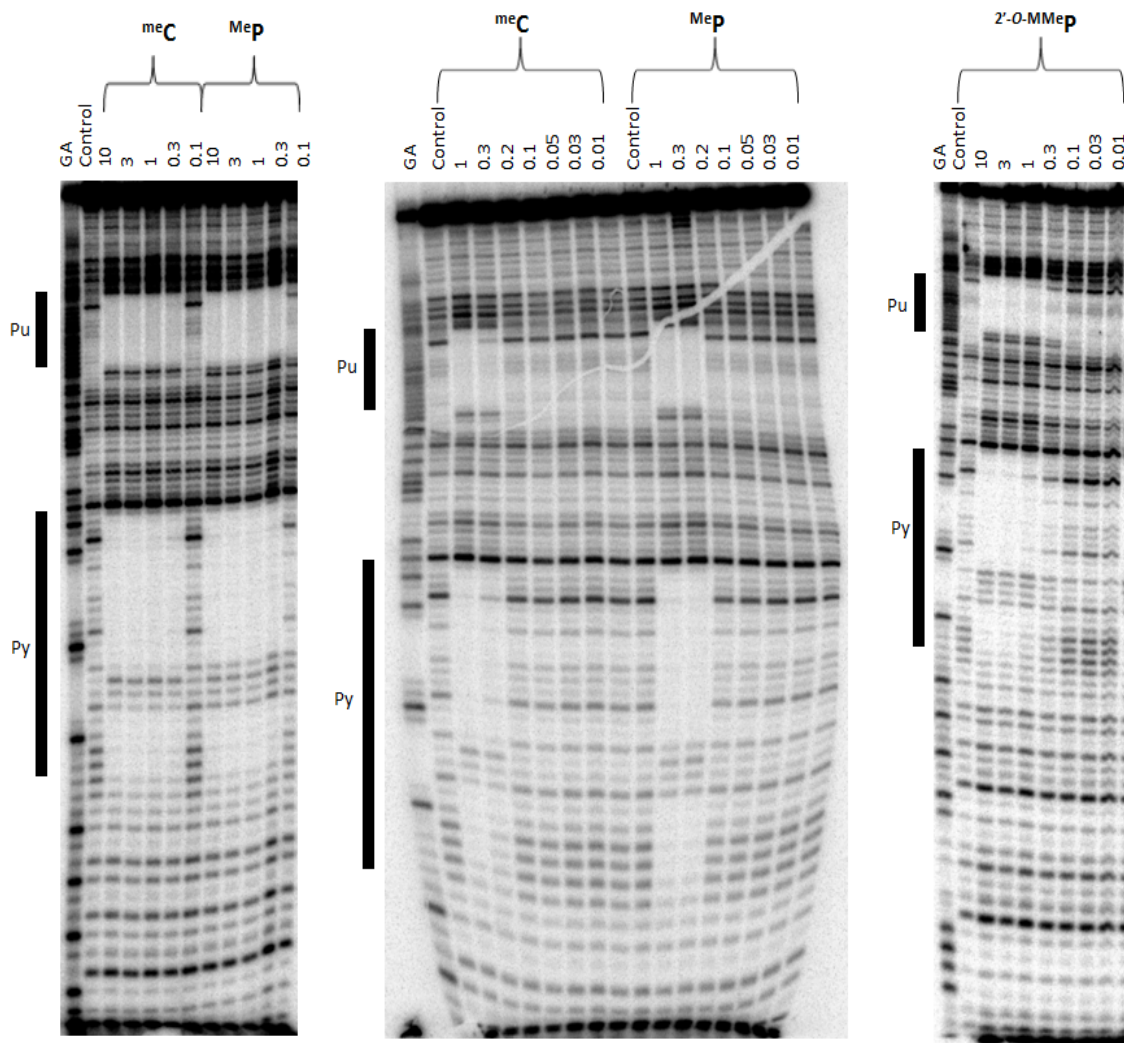


Fig. 3.20: DNase I cleavage patterns produced by **OL3/MeP**, **OL4/MeC** and **OL2/MeP_{OMe}** with the 15 base-pair oligopurine target site interrupted by a cytosine, that exists within the *slo-1 promoter* gene region. Oligonucleotide concentrations (μ M) are indicated at the top of each gel lane. The reaction was performed in 10 mM Tris-HCl, pH 7.0. Track labelled “GA” is the sequencing lane that is specific for purines (G + A). The solid bars indicate the location of the target site; ‘Pu’ indicates the purine-rich strand of the target site; ‘Py’ indicates the pyrimidine-rich strand.

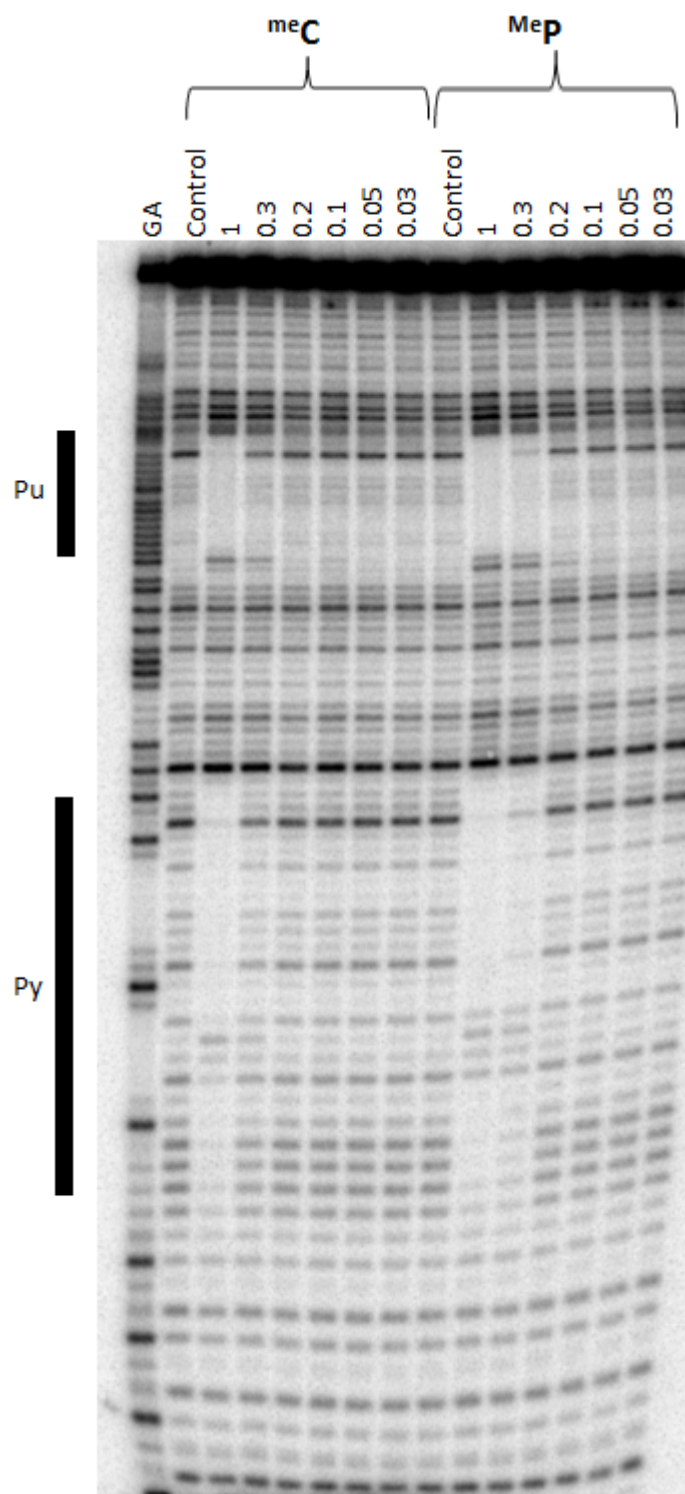


Fig. 3.21: DNase I cleavage patterns produced by **OL3^{MeP}** (right) and **OL4^{MeC}** (left) with the 15 base-pair oligopurine target site interrupted by a cytosine, that exists within the *slo-1* promoter gene region. Oligonucleotide concentrations (μM) are indicated at the top of each gel lane. The reaction was performed in 10 mM Tris-HCl, 50 mM NaCl pH 7.5. Track labelled “GA” is the sequencing lane that is specific for purines (G + A). The solid bars indicate the location of the target site; ‘Pu’ indicates the purine-rich strand of the target site; ‘Py’ indicates the pyrimidine-rich strand.

From *fig. 3.19* it is clear to see that both **OL3**/^{Me}**P** and **OL4**/^{Me}**C** generate good footprints at pH 5.0 as expected. For both TFOs, there exists enhanced cleavage at the 3'-(lower) end of the upper target site (purine-rich strand) at the triplex-duplex junction; this is a good indicator that the TFOs are binding (evidence for local oligonucleotide-induced DNA structural change).^{106,199} It cannot however, be determined for the TFO concentrations used at this pH, which gives the better protection of the target site from cleavage by DNase I.

At pH 7.0, the DNase I cleavage patterns (*fig. 3.20*) show that **OL3**/^{Me}**P** and **OL4**/^{Me}**C** prevent cleavage of the target site by DNase I efficiently, down to a concentration of 0.3 μ M. **OL3**/^{Me}**P** binds with a slightly higher affinity than **OL4**/^{Me}**C** at this concentration, while **OL2**/^{Me}**P**_{OMe} provides effective protection from cleavage down to a concentration of only 1.0 μ M. Hence, both **OL3**/^{Me}**P** and **OL4**/^{Me}**C** have a binding affinity, and in theory, an association rate constant (K_a , *this was not determined during this study*) that is approximately three-fold greater than that for **OL2**/^{Me}**P**_{OMe}. At pH 7.5 (*fig. 3.21*), **OL3**/^{Me}**P** provides the better protection from cleavage of the target site by DNase I, compared to **OL4**/^{Me}**C**. At a concentration of 0.3 μ M, **OL3**/^{Me}**P** displays high binding affinity to the target site, whereas **OL4**/^{Me}**C** has a similar binding affinity at a concentration (1.0 μ M) that is three times that of **OL3**/^{Me}**P**. Hence, the theoretical K_a for **OL3**/^{Me}**P** would be approximately three-fold greater than that for **OL4**/^{Me}**C**.

From these results, it can be stated that at physiological pH, **OL3**/^{Me}**P** provides the best protection of the target site from cleavage by DNase I and hence has the greatest binding affinity (theoretical K_a) when compared to **OL2**/^{Me}**P**_{OMe} and **OL4**/^{Me}**C**. Therefore, incorporation of ^{Me}**P** into triplex-forming oligonucleotides has a positive effect on the stability of the triplex which it forms with its target duplex, and hence an enhanced binding affinity at physiological pH.

DNase I footprinting to determine the effectiveness of TFO **OL5**/^{Me}**C**_{AE}, was not carried out as the 2'-O-aminoethyl group was found to be acetylated (discussed above). In the future a footprinting comparison needs to be made between ^{Me}**P** and ^{Me}**C**_{AE}, when the ^{Me}**C**_{AE} monomer has been synthesised without acetylation of the 2'-O-aminoethyl amine.

3.2.2 CONCLUSION

The utilisation of TFOs for genomic manipulation requires the development of oligonucleotides that can form stable triplexes in the nuclear environment. Base and sugar modifications have the potential to overcome the classical limitations imposed by physiological pH and Mg^{2+} concentration.²⁸⁶

In the experiments described above, the objective has been to identify which modified Cytosine-analogue when incorporated into a TFO of interest, is best able to enhance the stability of the triplex at physiological pH. From the footprinting results presented here, at pH 7.5, $^{\text{Me}}\text{P}$, when incorporated into the TFO of interest, enhances the stability of the triplex formed by three-fold relative to $^{\text{Me}}\text{C}$, this is due to its enhanced basicity. From both the UV melting and footprinting results, this TFO (**OL3**) has a higher affinity to double-stranded DNA over the pH range of 6.2-8.0 and a greater pH-independence in comparison with the other TFOs studied. Hence, $^{\text{Me}}\text{P}$ is an excellent replacement for $^{\text{Me}}\text{C}$ in recognition of double stranded DNA at physiological pH. However, due to the acetylation of the 2'-*O*-aminoethyl in $^{\text{Me}}\text{C}_{\text{AE}}$, this conclusion needs to be revisited when this TFO has been synthesised.

It is important to note at this stage that the effective delivery of oligonucleotides to their respective sites of action in the nucleus or cytoplasm has been an issue of concern.²⁹³ Although **OL2**/ $^{\text{Me}}\text{P}_{\text{OMe}}$ showed a lower stabilising effect on the triplex formed when compared to **OL4**/ $^{\text{Me}}\text{C}$ and **OL3**/ $^{\text{Me}}\text{P}$, nucleosides with 2'-modifications are known to have high resistance to degradation by nucleases *in vivo*, whereas 2'-deoxyribonucleosides are highly susceptible to this degradation, and are very unstable *in vivo*.^{294,295} Hence, although the *in vitro* studies show low stability by the triplex formed by **OL2**/ $^{\text{Me}}\text{P}_{\text{OMe}}$ at physiological pH when compared to those triplexes formed with **OL3**/ $^{\text{Me}}\text{P}$ and **OL4**/ $^{\text{Me}}\text{C}$, the T_m is reasonable considering the high resistance this oligonucleotide should experience against nuclease cleavage when used in biological systems.

It was essential that the above thermal stability studies were carried out; only then was it realised, by making a comparison to results reported by Cuenoud et al.²⁹¹ and Puri et al.,²⁸⁶ that the $^{\text{Me}}\text{C}_{\text{AE}}$ was acetylated.

The footprinting studies prove that by incorporating ^{Me}P into an oligo, the stability of the triplex formed is enhanced in comparison to an oligo containing either the naturally occurring ^{Me}C or the synthetic ^{Me}P_{OMe}, and hence this TFO efficiently protects the triplex from DNase I cleavage at physiological pH.

It is expected that once the synthesis of a suitable ^{Me}C_{AE} monomer has been established, the stability of the triplex formed with the respective TFOs will be profoundly enhanced. Also, due to the presence of the 2'-modification, *in vivo* experiments are expected to produce promising results due to the monomer's ability to inhibit TFO degradation by nuclease cleavage.

3.3 PSORALEN PHOTO- CROSSLINKING REACTIONS

3.3.1 *IN VITRO* STUDIES

3.3.1.1 UV-MELTING STUDIES - *Psoralen-conjugated TFO containing ^{Me}P*

Psoralen is a linear furocoumarin which reacts *via* a [2+2] cycloaddition, in concert with long wave UV light (UVA), to form photoadducts primarily with thymidines at 5'-TpA-3' sites in double stranded DNA. This reaction is highly regio- and stereospecific, forming an interstrand cross-link as the major product under the conditions employed in the experiments carried out.²⁸⁴

A 30-mer duplex (**7•8**), being made up of the complementary strands, **OL7** and **OL8** (*fig. 3.21* in which the 22-base pair TFO binding site is underlined), was used as the target for the 5'-end psoralen-linked TFO (**OL6**), and is a region of the *UNC-22* coding gene in *C. elegans*. 3-methyl-2-amino-5(2'-deoxy-β-D-ribofuranosyl)pyridine (^{Me}P) and Psoralen C6 (**Pso**) were incorporated into the TFO, **OL6**. Studies were carried out to compare using ultraviolet melting, the difference in the triplex thermal stability at different pH; pre and post UV photo-crosslinking of the Psoralen. A schematic representation of the sequences is outlined in *fig. 3.22*.

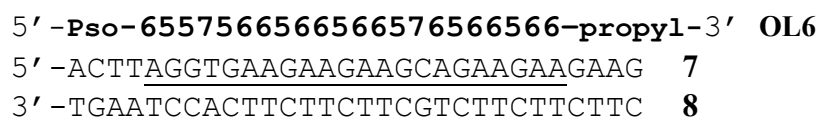


Fig. 3.22: Triplex UV melting experiment. TFO **OL6** shown in bold; **6** = T_{AE} ; **5** = ^{Me}P , **7** = S_{ME} and **Pso** = Psoralen (see page 280-281 for the respective structures of these base analogues). Oligonucleotides **7·8** form the duplex that contains the target site for the TFO, which is underlined.

The objective here was to demonstrate that by incorporating psoralen into the TFO **OL6** and carrying out photo-crosslinking reactions under long wavelength ultraviolet light, the thermal stability of the triplex would be enhanced significantly at pH values up to physiological pH, in comparison to that prior to photo cross-linking. Therefore the possibility of utilising the TFO *in vivo*.

T_m (°C)		
pH	Pre irradiation	Post irradiation
6.6	59.52	-
7.0	50.70	>70.0
7.5	41.66	67.47
8.0	31.93	58.24

Table 3.3: Thermal stabilities, T_m (°C) of the triplexes formed at varying pH, between the psoralen-linked oligonucleotide and duplex **7·8** ($T_m \sim 70$ °C), pre- and post- UV induced photo-crosslinking reactions.

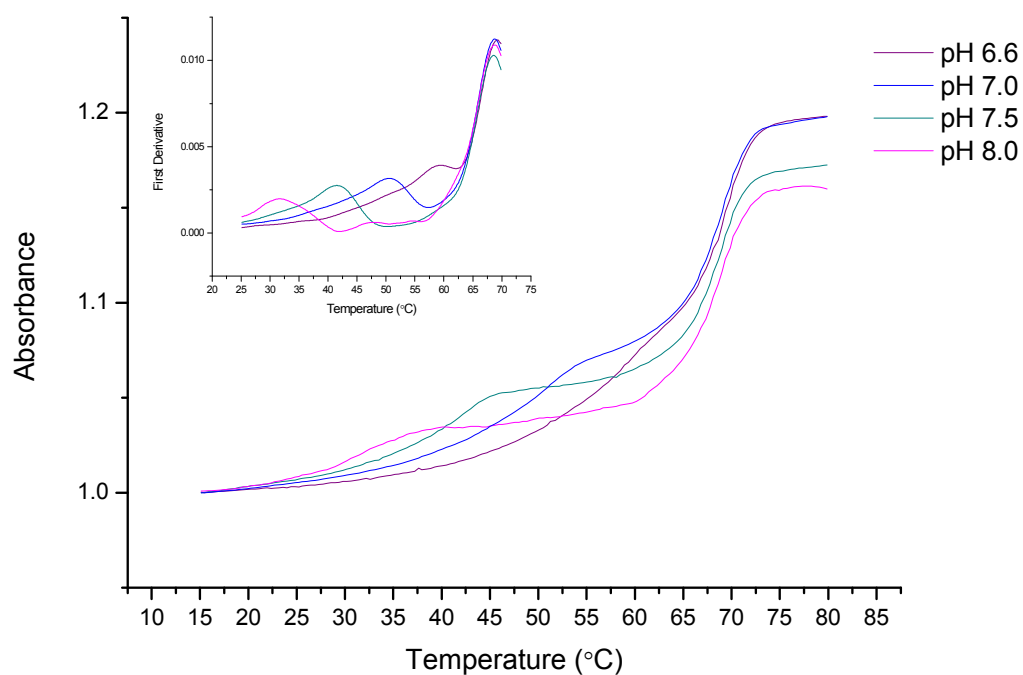


Fig. 3.23: UV-melting profiles and first derivatives for the complexes formed between TFO **OL6** and duplex **7·8** pre photo-crosslinking. The reactions were performed in 10 mM sodium phosphate, 200 mM NaCl, pH 6.6-8.0. The duplex concentration was 1 μ M and the TFO concentration was 3 μ M.

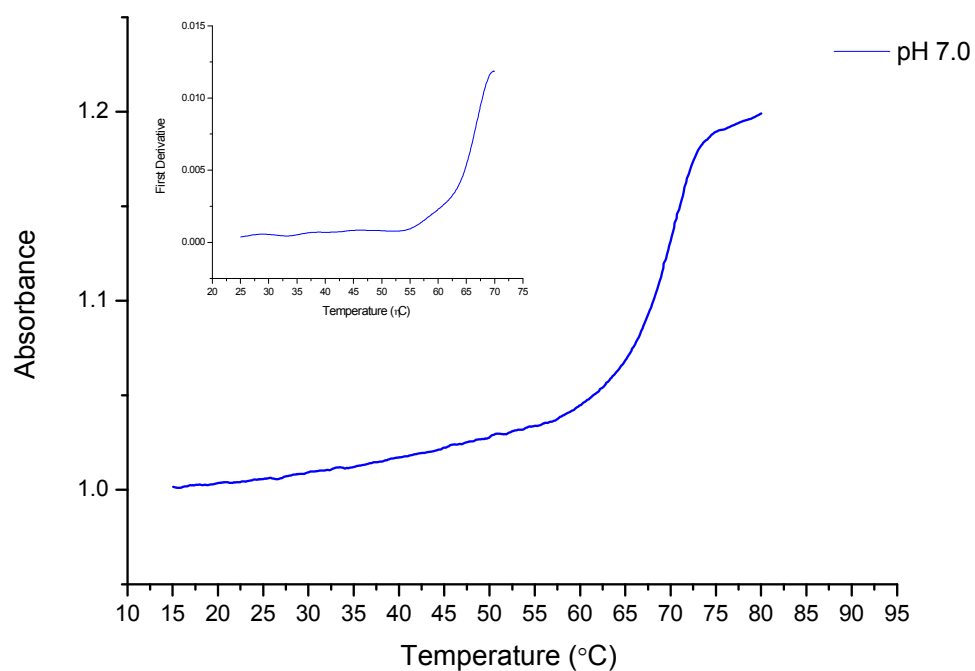


Fig. 3.24: UV-melting profiles and first derivatives for the complexes formed between TFO **OL6** and duplex **7·8** post photo-crosslinking. The reactions were performed in 10 mM sodium phosphate, 200 mM NaCl, pH 7.0. The duplex concentration was 1 μ M and the TFO concentration was 3 μ M.

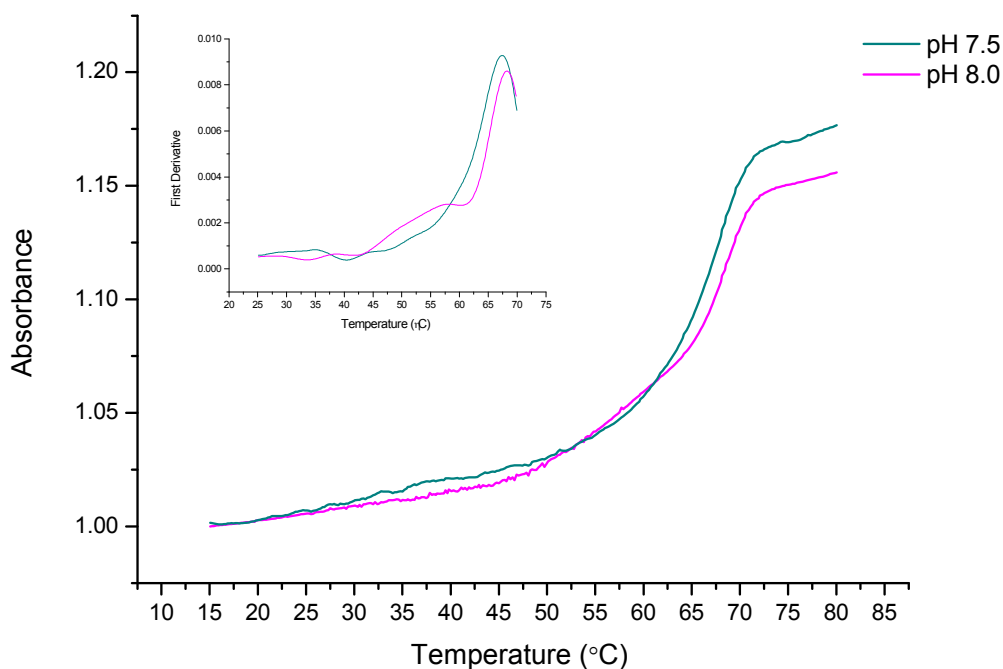


Fig. 3.25: UV-melting profiles and first derivatives for the complexes formed between TFO **OL6** and duplex **7·8** post photo-crosslinking. The reactions were performed in 10 mM sodium phosphate, 200 mM NaCl, pH 7.5-8.0. The duplex concentration was 1 μ M and the TFO concentration was 3 μ M.

The interaction of the Psoralen linked TFO **OL6** with the duplex DNA target **7·8** was initially studied by ultraviolet melting in the pH range 6.6-8.0 prior to photo-crosslinking. The results are displayed in the first column of *table. 3.3*, and in *fig. 3.23*. Significantly high thermal stabilities were observed within the pH range examined. The high T_m encountered can be attributed to 5'-Psoralen intercalating at the duplex-triplex junction, hence providing a powerful cooperative enhancement of binding to this target. As expected, the thermal stability of the triplex decreased with an increase in pH; $\sim 10^\circ\text{C}$ decrease in T_m was observed as pH was increased by 0.4-0.5 units.

The triplexes were then irradiated with UV light (365 nm for 20 mins) and subjected to further melting studies. In all cases, as expected, the melting curves were totally transformed (see *figs. 3.24-3.25*), in agreement with that reported by Li *et al.*²⁹⁶ After irradiation, at pH 7.0 and 7.5 there existed only one high-temperature melting transition, the transition owing to the dissociation of the TFO from the triplex structure had disappeared. The duplex and cross-linked TFO were melting at similar temperatures

~ 70 °C. Post irradiation at pH 8.0, two melting transitions were observed. The transition corresponding to the melting of the cross-linked triplex had shifted significantly to higher temperatures. At all pH studied, a shift in the triplex T_m of ~ + 26 °C was observed on forming the cross-linked triple helical complex from the pre-crosslinked triplex.

In conclusion, by using psoralen in combination with nucleotide analogues, and allowing triplex-directed double strand psoralen photoadduct formation, highly stable cross-linked structures are able to form under physiological conditions. The crosslinking studies substantiate the use of psoralen and UVA phototherapy *in vivo* as a conventional treatment for psoriasis. In addition, the studies demonstrate that diseases with a similar pathophysiology to psoriasis can be treated in a similar way.

3.3.1.2 FOOTPRINTING STUDIES- *Psoralen-conjugated TFO containing ^{Me}P*

DNase I footprinting was carried out to examine how the incorporation of ^{Me}P and psoralen within the TFO, **OL6** affects the recognition affinity and stability of the triple helix formed at a particular target site, at pH values up to physiological pH. The target site was a 22-base pair DNA fragment, part of the *UNC-22* coding gene within *C.elegans* (figure. 3.26).

OL6 5' - **Pso** - **6557566566566576566566** - propyl - 3'
 5' - GATCTACCCTGGAACAAAACTTAGGTGAAGAAGAAGCAGAAGAAGAAGAACT
 3' - CTAGATGGGACCTTGTTTTGAATCCACTTCTTCTTCGTCTTCTTCTTCTTGA

Fig. 3.26: Sequence of the 22 base-pair target site (underlined) within part of the *unc-22* coding gene which had been cloned into the pUC19 plasmid. TFO **OL6** shown in bold; **5** = ^{Me}P; **6** = T_{AE}; **7** = S_{ME} and **Pso** = **Psoralen** (see page 280-281 for the respective structures of these base analogues).

A restriction fragment (pyrimidine rich strand) labelled at the 3'-end with ³²P, containing the target site outlined in fig. 3.26, was incubated at room temperature overnight with increasing concentrations of the triple helix forming oligonucleotide **OL6** (0.03 μM-10 μM). DNase I generated reaction products were separated by

denaturing gel electrophoresis, and after fixing, drying, and storing overnight in a phosphorimager screen, the footprint was visualised by scanning.

From *fig. 3.27* it can be seen that protection of the target site from DNase I cleavage by **OL6** occurs at a concentration of 10 μM ; between 10-1.0 μM **OL6** has attenuated DNase I cleavage, it is slightly better at pH 7.0 than at pH 7.5. Therefore the triplex formed at pH 7.0 has a slightly higher stability than that at pH 7.5. The results demonstrate that the stability of the triplex decreases with increasing pH. At a 10 μM concentration of the TFO a stable triplex is formed at physiological pH, the TFO is able to protect the target site from DNase I cleavage.

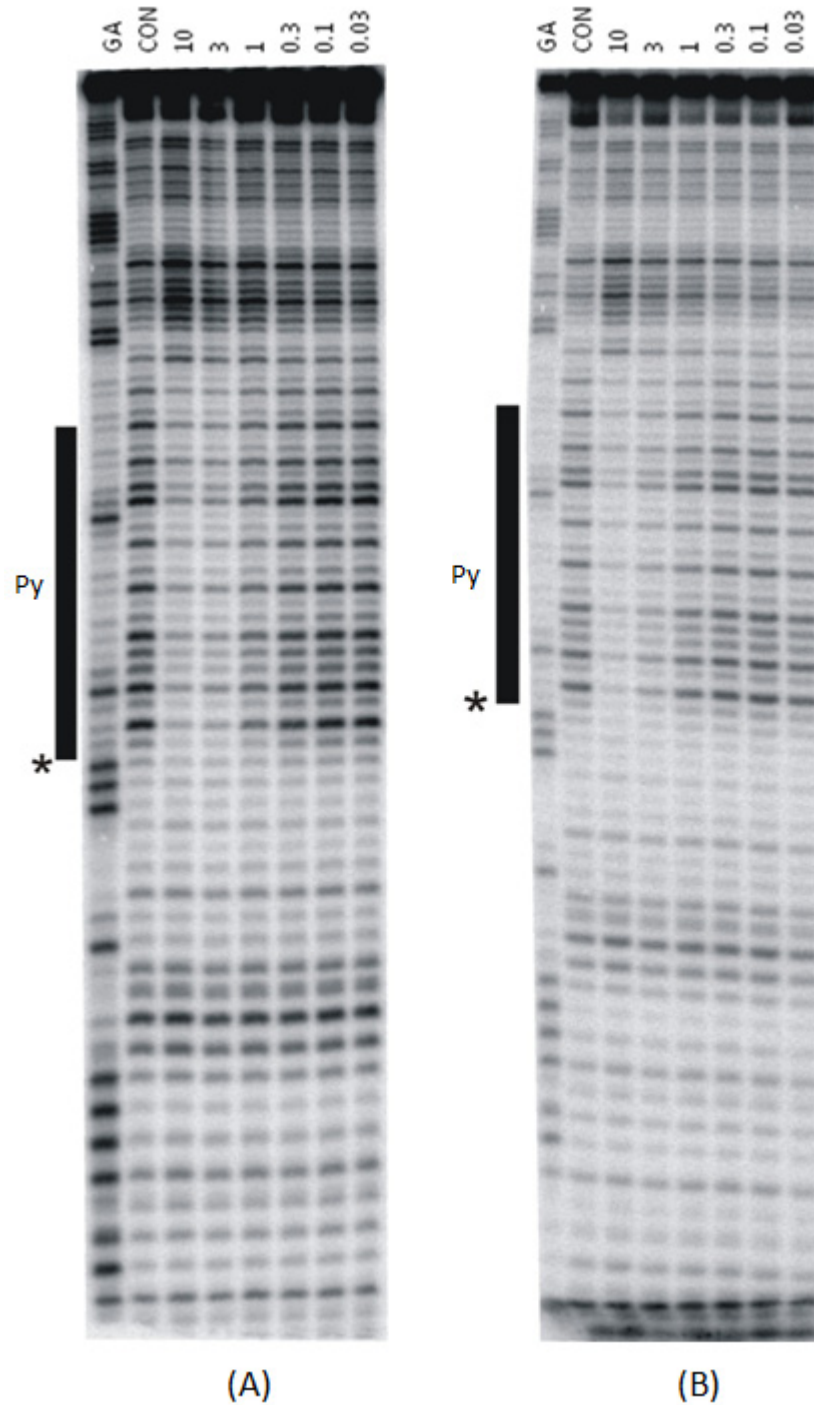


Fig. 3.27: DNase I cleavage patterns produced by the interaction of TFO **OL6** with the 22 base-pair target site within the *UNC-22* coding region in *C. elegans*; the pyrimidine strand is visualised here. Oligonucleotide concentrations (μM) are indicated at the top of each gel lane. The reactions were performed in: (A) 10 mM Tris-HCl, pH 7.0, containing 50 mM NaCl and (B) 10 mM Tris-HCl, 50 mM NaCl pH 7.5. Tracks labelled “GA” are sequencing lanes that are specific for purines (G + A); tracks labelled “CON” represent control where there was no TFO present. These were obtained without any UV-irradiation. The solid bars indicate the location of the pyrimidine-rich target site and the asterisk corresponds to the TpA step where the psoralen will bind.

3.3.1.3 GEL ELECTROPHORESIS

Gel-retardation analysis was utilised to determine the ability of the psoralen labelled TFO **OL6** to form double-stranded cross-links with its target duplex, a 22-base pair DNA fragment, part of the *UNC-22* coding sequence within *C.elegans* (fig. 3.26). One strand of the duplex was labelled at the 3'-end with [³²P] phosphate. The TFO, **OL6** was equilibrated with the duplex overnight at 20 °C before irradiating. After ultraviolet irradiation at 365 nm, the samples were denatured, and the reaction products analysed by gel electrophoresis (5 % polyacrylamide gel, fig. 3.28).

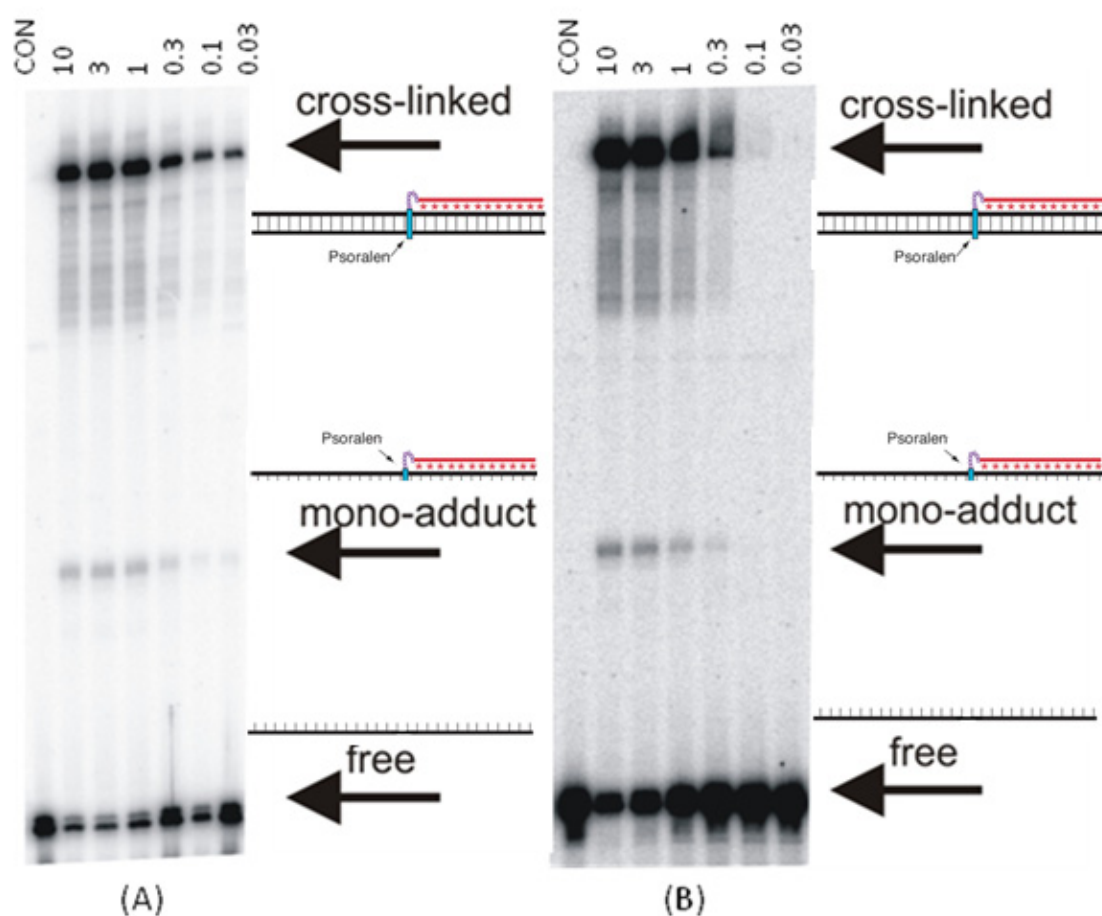


Fig. 3.28: Gel electrophoresis depiction of the UV-induced crosslinking of the psoralen-linked TFO **OL6** to the target duplex, the 22-base pair tract, part of the *unc-22* coding sequence in *C.elegans*. The reactions were performed in (A) 10 mM Tris-HCl, pH 7.0, containing 50 mM NaCl, and (B) 10 mM Tris-HCl, 50 mM NaCl pH 7.5. Oligonucleotide concentrations (μM) are indicated at the top of each gel lane. “free”, “mono-adduct”, and “cross-linked” indicate the presence of a single DNA strand, the formation of covalent attachment between the TFO **OL6** and one strand of the duplex, and the covalent attachment between the TFO **OL6** and both strands of the duplex, respectively; illustrated in the diagrams.

At both pH, a pattern of retarded bands is produced; the lower retarded species corresponds to the mono-adduct, while the upper (slower) retarded band corresponds to the bis adduct in which the TFO is linked to both strands of the duplex (*fig. 3.28*). The slower, bis-adduct is the predominant species after 20 mins UV irradiation, while a small amount of mono-adduct is still present. This mono-adduct could represent; (1) where the first reaction of psoralen with a thymidine in the target site involved the pyrone 3,4-double bond, this mono-adduct does not then go on to absorb light above 300 nm and cannot be converted to the bis adduct; or (2) it could represent incomplete cross linking where only the initial cross linking of the furan moiety of psoralen to the pyrimidine strand has occurred, which would then if the reaction went to completion, be followed by cross-linking of the pyrone moiety to the purine strand. The bis-adduct represents where the initial reaction of psoralen involved the 4',5'-double bond of the furan ring to form the mono-adduct with the 5,6-double bond of the thymidine in the purine strand of the duplex target site. This then goes on further to absorb light above 310 nm, in order to engage in the second reaction involving the 3,4-double bond on the pyrone side with the 5,6 double bond of thymidine in the pyrimidine strand of the duplex target site.^{296,297}

The UV-induced crosslinking is more efficient at pH 7.0 than at pH 7.5. At pH 7.0, some cross-linking is evident down to a concentration of 0.03 μ M. At pH 7.5 no cross-links are produced below a concentration of 0.3 μ M. Hence, crosslinking at pH 7.0 is possible at a concentration ten-fold less than that at pH 7.5.

From these results it can be stated that, incorporating ^{Me}P and psoralen into TFOs, enables high affinity recognition of the TFO to the target site, hence formation of stable triplexes at physiological pH. The additional feature of crosslinking two strands of DNA at specific sites *via* UV irradiation enables the formation of covalent attachments between the psoralen, TFO and duplex target site. Biological processes that require an opening of the double helix such as replication and transcription can therefore be blocked. Inside cells it is likely that these cross-links will be recognised by repair enzymes. This could lead to mutations at the specific sites where the oligonucleotide has targeted the cross-linking reaction.²⁹⁷ Hence, psoralen-containing TFO conjugates have the potential to be good probes for demonstrating the formation of triple helices

within living cells. This was investigated *via* the *in vivo* methods described below in *C. elegans*.

3.3.2 *IN VIVO* STUDIES

3.3.2.1 SOAKING

Soaking²¹⁰ was administered as the method of delivery of TFO, **OL6** to the *C. elegans* by external application (worms are soaked in a solution of the TFO **OL6**, this is the method of delivery). It has the disadvantage that it may have low efficiency in terms of penetrance of effects, but has the tremendous advantage that it is amenable to large, automated genome wide-screens. It was thought to be the most suitable method to test for the effects of TFO mediated regulation of gene expression in adult animals.

Experiments were carried out, with and without the aid of nanoparticles, *in vivo*-jetPEITM; a delivery reagent used to aid the delivery and improve the uptake of **OL6** by the cells. *In vivo*-jetPEITM was developed to deliver DNA and oligonucleotides in order to mediate gene expression in various tissues upon *in vivo* administration. It does not induce any significant pro-inflammatory response after application. Unfortunately, in both cases, whether or not the delivery reagent was administered, no effect was observed.

UNC-22 provides a sensitive and specific assay for genetic interference, as it is the only gene in *C. elegans* genome that can mutate by loss-of-function, to give a twitching phenotype.²⁹⁸ Hence, it was expected that **OL6** would induce a decrease in *UNC-22* gene expression and therefore produce twitching phenotypes in the *C. elegans* that were soaked. However this was not observed.

An alternative delivery route should be attempted in the future. It is possible that injection into syncytial tissue (*C. elegans* gonad) for incorporation into germ cells could provide efficient uptake of the TFO. When generating transgenic animals, this is the route of access for plasmid DNA. It is anticipated that TFOs will gain intracellular access to the germ cells when delivered by this route. This method of delivery however, has the disadvantage that the effect of the TFO is likely to be diluted out following cell

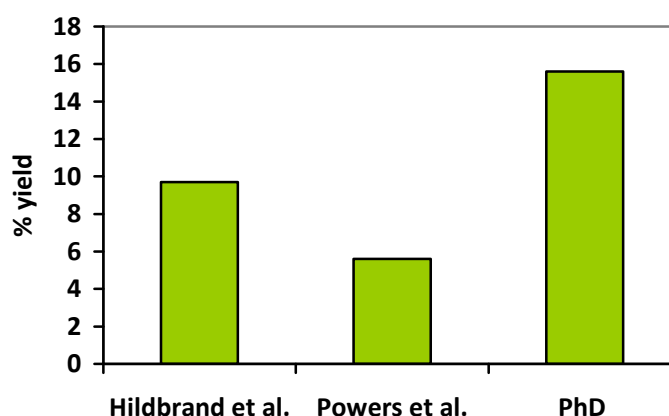
division, so the highest penetrance of effect might be expected only at the earliest stages of development. Alternatively, the TFOs can be administered by injection into the intestine of *C. elegans*. This method has the unfortunate disadvantage that it is relatively slow and not suitable for targeting large numbers of genes. These two delivery methods are worth attempting in the future. If proved successful, this research can progress further towards creating therapeutics for diseases caused by point mutations.

CHAPTER 4

Conclusions

4.1 THE ACHIEVEMENTS OF THIS RESEARCH

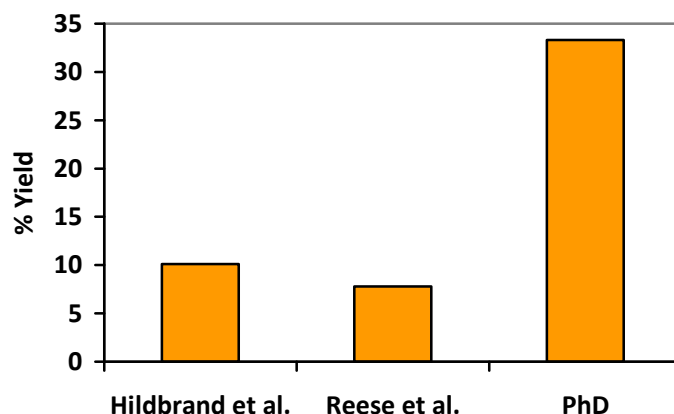
The initial aims set out for this PhD research have all been met successfully. A novel, concise synthetic route to the ^{Me}P monomer has been developed. The overall yield obtained for this monomer (starting from the coupling reaction of the carbohydrate moiety with the aglycon) is significantly higher than those reported (15.5 % *vs.* 9.7 %¹⁵⁸ and 5.6 %²⁰¹; see graph. 1 below). The number of steps required to obtain the final phosphoramidite is less than what has been reported (10 *vs.* 14¹⁵⁸ and 15²⁰¹).



Graph. 1: A comparison between the yields obtained for the ^{Me}P monomer by others, and that achieved during this PhD research.

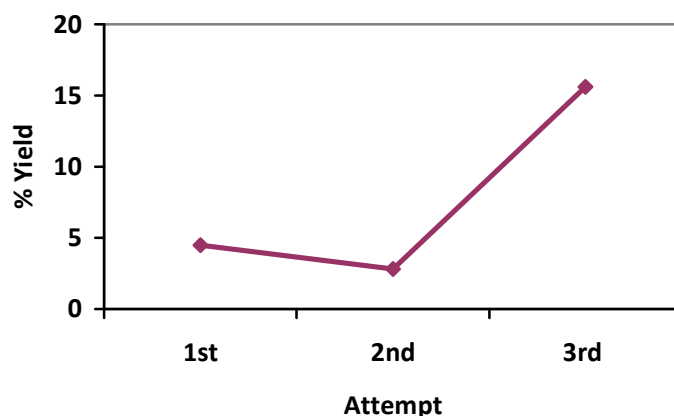
In addition, the overall yield to the free nucleoside is significantly higher than that reported (33.3 % *vs.* 10.1 %¹⁵⁸ and 7.8 %²⁰²; see graph. 2 below). The synthetic scheme developed here to obtain the ^{Me}P monomer is efficient, and it makes a good contribution towards nucleic acids chemistry. Others needing to develop this cytidine analogue for use in biological studies/applications can use the route developed here.

However, during the second attempt at optimising the synthetic route, there was a drop in the overall yield of the final monomer (graph. 3). This was due to the low yielding tritylation (40 %) and phosphitylation (34 %) reactions. All other steps for this synthetic scheme afforded yields that are ≥ 60 % (the majority of which are ≥ 80 %).



Graph. 2: A comparison between the yields obtained for the free ^{Me}P nucleoside by others, and that achieved during this PhD research.

In conclusion, the exocyclic amino moiety of ^{Me}P needs to be blocked prior to the tritylation reaction in order to obtain a tritylation yield of ≥ 60 %. The tritylation step in the first attempt (see *scheme. 2.38*, page 129) was ~ 62 %; the exocyclic amino moiety in the precursor for this step was protected. In the second attempt (see *scheme. 2.45*, page 135), tritylation was carried out on the free nucleoside, **1** β . Although the exocyclic amino moiety did not partake in the tritylation reaction, it may interfere, causing a reduction in this yield. Hildbrand *et al.*¹⁵⁸ obtained a tritylation yield of 74 %; the precursor to this reaction was protected at the exocyclic amino moiety with phenoxyacetyl (see *scheme. 2.15*, page 105).



Graph. 3: The progression of the overall yield of ^{Me}P after optimisations of the synthetic route.

The greatest overall yield for ^{Me}P was achieved during the third attempt (see graph. 3). However, there is still room for improvement. The phosphitylation reaction yield can be improved using reagents with a greater purity. In addition, the mistunobu reaction needs further optimisation; the $\beta:\alpha$ ratio for this step was 3:2. As discussed in Chapter 2, methods that can increase the ratio of β -glycosylation need to be investigated, and this is an area that this research can be extended. Increasing the $\beta:\alpha$ ratio can significantly increase the overall yield of the ^{Me}P monomer. As mentioned in Chapter 2, using bulkier protecting groups, such as *tert*-butyldiphenylsilyl (TBDPS), at the 3'- and 5'-positions of the 2'-deoxy- β -D-ribose ring, may help to increase β -glycosylation. During the coupling reaction of lactol, **14** with the protected base, **86**, the TBDPS group at the 3' position should increase steric hindrance at the *exo* face of the sugar ring, and can therefore potentially induce an increase in the diastereoselectivity towards the epimer that then cyclises under mistunobu conditions to give the β -product. Alternatively, attempting *scheme. 2.31* in chapter 2 (page 122), is another way that could increase β -glycosylation. This would involve oxidation of lactol, **14** to the 2'-deoxy lactone, **108**, and following coupling, the stereoselective reduction of the hydroxyl ketone, **109** with L-Selectride to the diols, **103 RS**. As a result of its steric nature, L-Selectride (a hydride reagent), can stereoselectively reduce carbonyl groups. The scheme would therefore be increased by two steps, which would take the total to 12 steps. However, if these steps are high yielding, and the percentage of β -glycosylation is increased, it would be worthwhile.

In addition to the development of a concise, high yielding route to the monomer ^{Me}P, *in vitro* and *in vivo* studies of ^{Me}P containing pyrimidine motif TFOs have been carried out. No footprints were detected for TFO **OL1** at physiological pH. This was a result of the dimethylformamidinium group not being suitable for the protection of the exocyclic amino moiety. This protecting group converts to the inert formyl group during conditions of standard oligonucleotide deprotection.

^{Me}P can be used in place of cytidine in TFOs in the pyrimidine motif to give stable triplexes. Footprinting and UV melting studies comparing the ^{Me}P containing pyrimidine motif TFO, **OL3** with other cytidine analogue substituted pyrimidine motif TFOs (^{Me}C/**OL4** and ^{Me}P_{OMe}/**OL2**), demonstrated that for achieving stable triplexes at physiological pH, ^{Me}P is superior to both ^{Me}C and ^{Me}P_{OMe}. The poor biophysical results

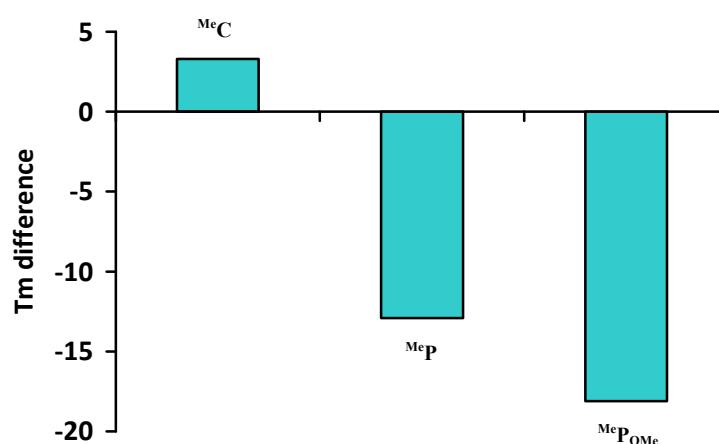
displayed by the $^{\text{Me}}\text{C}_{\text{AE}}$ incorporated TFO, **OL5**, indicated the need for further optimisation of the synthetic route to this monomer. In particular, it was discovered that during the Staudinger reaction, an unwanted acetyl migration was occurring from the exocyclic amino moiety on the base of $^{\text{Me}}\text{C}_{\text{AE}}$, to the 2'-*O*-aminoethyl moiety. Hence, the comparison of $^{\text{Me}}\text{P}$ vs $^{\text{Me}}\text{C}_{\text{AE}}$ biophysical results needs revision. The synthetic scheme for $^{\text{Me}}\text{C}_{\text{AE}}$ has been optimised (by Miss M. Shelbourne) to avoid the unwanted side reaction, hence for future studies the same TFO sequence incorporating the newly synthesised $^{\text{Me}}\text{C}_{\text{AE}}$ **OL5** needs to be compared to **OL3**.

The simultaneous incorporation of $^{\text{Me}}\text{P}$ and psoralen into a pyrimidine motif TFO, **OL6** enabled the formation of stable triplexes to its target site at physiological pH. Following UV irradiation, the TFO demonstrated the ability to form crosslinked triplexes at its target site with high stability at physiological pH. *In vivo* studies of this TFO did not unfortunately demonstrate the formation of triplexes with a target site in *C. elegans*. This result however, does not imply that the TFO is not able to form stable triplexes *in vivo*, more so, it could be that the delivery method that was not effective. "Soaking" was the method of delivery of the TFO, and due to time constraints other delivery methods (*via* intramuscular, subcutaneous, intravenous, and intraperitoneal injection) were not attempted. Particle bombardment using gold particles has been shown to be able to deliver genes successfully into cells,²⁹⁹ and this is a method that should be explored for the delivery of TFO **OL6**. In addition, using electric pulses to introduce this TFO into cells may also be a successful approach. Electric pulses enable cell membranes to be transiently permeable.³⁰⁰ These are areas that need to be investigated further (see Chapter 3 for additional delivery method suggestions).

4.2 RESEARCH CONTINUED

Following on from the biophysical studies carried out to compare cytidine analogues, during this PhD research, further studies have been carried out, and reported after the duration of this PhD, by Mr. C. Lou²⁸² in this research group. A comparison between oligos containing dC, $^{\text{Me}}\text{P}$, and $^{\text{Me}}\text{P}_{\text{OMe}}$ in terms of their selectivity towards duplex DNA over single stranded DNA, and their resistance to nuclease degradation has been made. Overall, incorporating four additions of the cytidine analogues, $^{\text{Me}}\text{P}$, and $^{\text{Me}}\text{P}_{\text{OMe}}$ separately into two different polyT TFOs, both TFOs had greater selectivity towards duplex DNA than single stranded DNA. This was demonstrated by the negative ΔT_m

(the difference between that particular TFO and the TFO containing dC) displayed on graph. 4. In particular, the TFO containing $^{\text{Me}}\text{P}_{\text{OMe}}$ had greater selectivity towards double stranded DNA than the TFO containing $^{\text{Me}}\text{P}$. When $^{\text{Me}}\text{C}$ was incorporated into a TFO of this nature, the TFO had preference to the complementary antiparallel single strand of DNA than the duplex DNA (Graph. 4). This result demonstrates that *in vivo*, pyrimidine motif TFOs with a uniform ribose backbone containing either $^{\text{Me}}\text{P}$ or $^{\text{Me}}\text{P}_{\text{OMe}}$, will not encounter significant binding to single stranded DNA.



Graph. 4: The difference in UV melting temperatures (T_m) between the duplex formed from the TFO containing the respective cytidine analogue (indicated on graph) and its antiparallel complementary single DNA strand, and the duplex formed from the same single stranded DNA with the TFO containing dC.²⁸²

In addition, a homopyrimidine TFO containing $^{\text{Me}}\text{P}_{\text{OMe}}$ encounters greater resistance to degradation by serum nuclease than when $^{\text{Me}}\text{P}$ was incorporated. After 24 hours the TFO containing $^{\text{Me}}\text{P}$ was completely degraded while the TFO containing $^{\text{Me}}\text{P}_{\text{OMe}}$ was relatively intact.²⁸²

4.3 APPLICATIONS FOR GENE THERAPY?

From the biophysical comparison of the cytidine analogues described within this thesis, it can be hypothesised that $^{\text{Me}}\text{P}$ incorporated TFOs will be successful *in vivo* for future gene silencing therapeutics. Gene inhibition can be achieved by the highly selective and stable triplex formation at a promoter site, this has the potential to block the binding of various transcription factors, hence inhibiting transcription.⁷⁵ However it is well known that 2'-deoxyribose containing oligonucleotides are subject to rapid degradation by

nucleases,^{293-295,301} hence the likeliness of the oligo to be at its target site for a substantial amount of time to enable an effect is quite low. Therefore, further modifications to the ^{Me}P monomer, such as to the phosphodiester backbone (see Chapter 1; section 1.6.4.1.3) may be required to enable greater resistance to degradation by nucleases, and allow for wider applications *in vivo*.

4.4 BIOLOGICAL BARRIERS TO GENE THERAPY^{293,302}

For antigene TFOs to be successful for therapeutic value, they must be able to enter cells with minimal toxicity.³⁰² In addition, they must be able to overcome biological barriers, antigene TFOs are presented with many biological barriers that stand between their initial administration and their site of action. As discussed above, nuclease activity in the plasma and tissues, by exonucleases and endonucleases cause degradation to TFOs that are not sufficiently modified. TFOs can also encounter uptake by the kidney or liver which can cause rapid excretion. In addition to this, the reticuloendothelial system can clear the body of any foreign TFOs. Endothelial cells control the integrity of the vasculature and these are also a barrier to TFOs reaching their target site.²⁹³ Progress is being made to overcome these biological barriers to gene therapy, so that in the future these therapeutics can be widely available for the management of diseases at the transcriptional level. This will affect the medical industry in a positive manner, and change the way that healthcare is practised.

4.5 MOVING FORWARD^{122,125}

As discussed in chapter 1, PNA is an oligonucleotide mimic where the deoxyribose phosphate backbone of DNA has been replaced by a backbone composed of 2-aminoethylglycine units. This is an area that can be expanded for the 3-methyl-2-aminopyridine base. Developing a synthetic route to the PNA equivalent of ^{Me}P, and carrying out biophysical analyses on triplexes formed with respective DNA targets, would be an area where this PhD research could be extended. Pyrimidine motif PNA oligos containing natural nucleosides are able to form highly stable (PNA)₂·DNA triplexes that are much more stable than the corresponding DNA·DNA duplexes.¹²² From the biophysical results obtained during this PhD research, in terms of providing triplex stability, the following trend is observed: ^{Me}P > ^{Me}C > dC, when these nucleosides are incorporated into TFOs. Hence, pyrimidine motif PNA oligos containing ^{Me}P would be expected to create (PNA)₂·DNA triplexes with even greater

stability than those containing natural nucleosides. It would be interesting to carry out the biophysical analyses, and if proved successful, test the PNA TFOs *in vivo*. They may be better at overcoming the biological barriers experienced by DNA and RNA TFOs.

CHAPTER 5

Experimental

5. EXPERIMENTAL

5.1 SYNTHESIS

5.1.1 GENERAL EXPERIMENTAL

5.1.1.1 REAGENTS AND SOLVENTS

All reagents were purchased from Aldrich, Avocado, Fluka and TCI Europe and used without purification with the exception of the following solvents, which were purified by distillation: Ethanol/Methanol (over iodine and magnesium); DCM/TEA/DIPEA (over calcium hydride); Pyridine (over potassium hydroxide); THF (over sodium wire and benzophenone). All reactions were carried out under an argon atmosphere using oven-dried glassware with purified and distilled solvents. Separation and purification by ion exchange was performed using Dowex resin 50WX2-200 (H⁺).

5.1.1.2 CHROMATOGRAPHY

Column chromatography was carried out under pressure using Fisher scientific DAVISIL 60A (35-70 micron) silica.

Thin layer chromatography was performed using Merck Kieselgel 60 F₂₄ (0.22mm thickness, aluminium backed). Compounds were visualised by irradiation at 365/254 nm, and/or staining with the following stains:

A' - Anisaldehyde (*p*-anisaldehyde 13.8 mL, glacial acetic acid 5.7 mL, conc. Sulphuric acid 18.75 mL, 95 % ethanol 507 mL (9.3:3.8:12.5:338 v/v)). Compounds containing a sugar component would produce blue/black colour after staining and heating. Compounds containing DMT component produced a strong orange colour after staining and heating.

B' – Ninhydrin (Ninhydrin 0.3 g, n-butanol 97 mL, glacial acetic acid 3 mL). Compounds containing a primary amine group would produce a pink colour after staining and heating.

The solvent system for TLC referred to as H-1 is the following; EA: Acetone: EtOH: H₂O 4:1:1:1.

5.1.1.3 MELTING POINT ANALYSIS

Melting points were measured using a Gallenkamp electrothermal device and values have been quoted as a range of temperature (°C).

5.1.1.4 INFRARED SPECTROSCOPY

Infrared spectra were recorded on a BIORAD FT-IR using a Golden Gate adapter and BIORAD WIN-IR software or a Satellite FT-IR using a Golden Gate adapter and WIN FIRST-lite software. Absorptions are described as strong (s), medium (m), broad (b) or weak (w).

5.1.1.5 MASS SPECTROMETRY

Low-resolution mass spectra were recorded using electrospray technique on a Fisons VG platform instrument or a Waters ZMD quadrupole mass spectrometer in acetonitrile, water or methanol (HPLC grade). High-resolution mass spectra were recorded in acetonitrile, water or methanol (HPLC grade) using electrospray ionization on a Bruker APEX III FT-ICR mass spectrometer.

5.1.1.6 NMR SPECTROSCOPY

¹H NMR spectra were measured at 300 MHz on a Bruker AC300 spectrometer and 400 MHz on a Bruker DPX400 spectrometer. ¹³C NMR spectra were measured at 75 MHz and 100 MHz on the same spectrometers respectively. Chemical shifts are given in ppm relative to tetramethylsilane, and *J* values are given in Hz and are correct to within 0.5 Hz. All spectra were internally referenced to the appropriate residual undeuterated solvent signal.³⁰³ Assignment was also aided by COSY (¹H-¹H), HMQC (¹H-¹³C) and HMBC experiments. ³¹P NMR spectra were recorded on a Bruker AV300 spectrometer at 121 MHz and were externally referenced to 85 % phosphoric acid in D₂O. Multiplicities are described as singlet (s), doublet (d), triplet (t), quartet (q), multiplet (m), double doublet (dd), double triplet (dt), triple doublet (td), double double doublet (ddd), broad (br).

5.2 COMPOUNDS SYNTHESISED

No.	Compound	Page
1 α,β	2-Amino-3-methyl-5-(2'-deoxy-D-ribofuranosyl)pyridine	243
1 α	2-Amino-3-methyl-5-(2'-deoxy- α -D-ribofuranosyl)-pyridine	256
1 β	2-Amino-3-methyl-5-(2'-deoxy- β -D-ribofuranosyl)pyridine	245, 247, 261
14	3',5'-di- <i>O</i> -benzyl-2'-deoxy-D-ribo-1,4-lactol	228
64	2',3',5'-Tri- <i>O</i> -benzyl-D-ribo-1,4-lactone	202
68 β	2-Amino-3-methyl-5-(2',3',5'-tri- <i>O</i> -benzyl- β -D-ribofuranosyl)pyridine	207
69 β	3-Methyl-2-[<i>N</i> -(phenoxyacetyl)amino]-5-(2',3',5'-tri- <i>O</i> -benzyl- β -D-ribofuranosyl)pyridine	217
70 β	3-Methyl-2-[<i>N</i> -(phenoxyacetyl)amino]-5-(β -D-ribofuranosyl)pyridine	218
71 β	3-Methyl-2-[<i>N</i> -(phenoxyacetyl)amino]-5-[3',5'- <i>O</i> -(1,1,3,3,-tetraisopropyl disiloxane-1,3-diyl)- β -D-ribofuranosyl]pyridine	222
72 β	3-Methyl-2-[<i>N</i> -(phenoxyacetyl)amino]-5-[3',5'- <i>O</i> -(1,1,3,3,-tetraisopropyl disiloxane-1,3-diyl)-2'- <i>O</i> -(imidazolyl thiocarbonyl)- β -D-ribofuranosyl]pyridine	224
73 β	3-Methyl-2-[<i>N</i> -(phenoxyacetyl)amino]-5-[3',5'- <i>O</i> -(1,1,3,3,-tetraisopropyl disiloxane-1,3-diyl)-2'-deoxy- β -D-ribofuranosyl]pyridine	225
85	Methyl- 2',3',5'-tri- <i>O</i> -benzyl-D-ribose	199
86	5-Bromo-2-[<i>N,N</i> -bis(4-methoxybenzyl)]amino-3-methylpyridine	203
90	2',3',5'-Tri- <i>O</i> -benzyl-D-ribose	201
91	2-[<i>N,N</i> -bis(4-methoxybenzyl)]amino-3-methyl-5-(1'-hydroxyl-2',3',5'-tri- <i>O</i> -benzyl-D-ribofuranosyl)pyridine	205
92 β	2-[<i>N,N</i> -bis(4-methoxybenzyl)]amino-3-methyl-5-(2',3',5'-tri- <i>O</i> -benzyl- β -D-ribofuranosyl)pyridine	206
93 β	3-Methyl-2-trifluoroacetamide-5-(2',3',5'-tri- <i>O</i> -benzyl- β -D-ribofuranosyl)pyridine	208
94 β	2-Amino-3-methyl-5-(β -D-ribofuranosyl)pyridine	209, 218

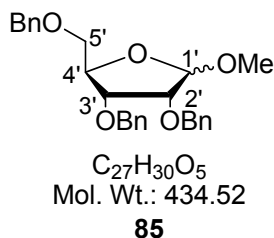
95 β	2- <i>N,N</i> -[(dimethylamino)methylidene]-3-methyl-5-(2',3',5'-tri- <i>O</i> -benzyl-β-D-ribofuranosyl)pyridine	210
96 β	2- <i>N,N</i> -[(dimethylamino)methylidene]-3-methyl-5-(β-D-ribofuranosyl)pyridine	211
97 β	2- <i>N,N</i> -[(dimethylamino)methylidene]-3-methyl-5-(3',5'- <i>O</i> -[1,1,3,3-tetraisopropylidisiloxane-1,3-diyl]-β-D-ribofuranosyl)pyridine	213
98 β	2- <i>N,N</i> -[(dimethylamino)methylidene]-3-methyl-5-[3',5'- <i>O</i> -(1,1,3,3-tetraisopropylidisiloxane-1,3-diyl)-2'- <i>O</i> -(imidazolylthiocarbonyl)-β-D-ribofuranosyl]pyridine	215
99 β	2- <i>N,N</i> -[(dimethylamino)methylidene]-3-methyl-5-[3',5'- <i>O</i> -(1,1,3,3-tetraisopropylidisiloxane-1,3-diyl)-2'-deoxy-β-D-ribofuranosyl]pyridine	216
100 β	2-Amino-3-Methyl-5-[3',5'- <i>O</i> -(1,1,3,3-tetraisopropylidisiloxane-1,3-diyl)-β-D-ribofuranosyl]pyridine	220
101	Pentafluorophenyl 2-phenoxyethanote	221
102	Methyl-3',5'-di- <i>O</i> -benzyl-2'-deoxy-D-ribose	226
103 RS	2-[<i>N,N</i> -bis(4-methoxybenzyl)]amino-3-methyl-5-(3',5'-di- <i>O</i> -benzyl-2'-deoxy-D-ribofuranosyl)pyridine	229
104 α,β	2-[<i>N,N</i> -bis(4-methoxybenzyl)]amino-3-methyl-5-(3',5'-di- <i>O</i> -benzyl-2'-deoxy-D-ribofuranosyl)pyridine	232
106 α,β	2-[<i>N</i> -(4-methoxybenzyl)]amino-3-methyl-5-(3',5'-di- <i>O</i> -benzyl-2'-deoxy-D-ribofuranosyl)pyridine	236
107 α,β	2-Amino-3-methyl-5-(3',5'-di- <i>O</i> -benzyl-2'-deoxy-D-ribofuranosyl)pyridine	237
107 α	2-Amino-3-methyl-5-(3',5'-di- <i>O</i> -benzyl-2'-deoxy-α-D-ribofuranosyl)-pyridine	254
107 β	2-Amino-3-methyl-5-(3',5'-di- <i>O</i> -benzyl-2'-deoxy-β-D-ribofuranosyl)pyridine	239
114 α,β	3-Methyl-2-trifluoroacetamide-5-(3',5'-di- <i>O</i> -benzyl-2'-deoxy-D-ribofuranosyl)pyridine	241
114 α	3-Methyl-2-trifluoroacetamide-5-(3',5'-di- <i>O</i> -benzyl-2'-deoxy-α-D-ribofuranosyl)-pyridine	256

114 β	3-Methyl-2-trifluoroacetamide-5-(3',5'-di- <i>O</i> -benzyl-2'-deoxy-β-D-ribofuranosyl)-pyridine	242
115 α,β	2- <i>N,N</i> -[(dimethylamino)methylidene]-3-methyl-5-(3',5'-di- <i>O</i> -benzyl-2'-deoxy-D-ribofuranosyl)pyridine	243
116	2-Amino-3-methyl-5-(1',2'-deoxy-D-ribofuranosyl)pyridine	245
117 β	2-Amino-3-methyl-5-(3'- <i>O</i> -benzyl-2'-deoxy-β-D-ribofuranosyl)pyridine	247
118 α	2- <i>N,N</i> -[(dimethylamino)methylidene]-3-methyl-5-(2'-deoxy-α-D-ribofuranosyl)pyridine	252
118 β	2- <i>N,N</i> -[(dimethylamino)methylidene]-3-methyl-5-(2'-deoxy-β-D-ribofuranosyl)pyridine	248
119 β	2- <i>N,N</i> -[(dimethylamino)methylidene]-3-methyl-5-(5'- <i>O</i> -(4,4'-dimethoxytrityl)-2'-deoxy-β-D-ribofuranosyl)pyridine	249
120 β	2- <i>N,N</i> -[(dimethylamino)methylidene]-3-methyl-5-(5'- <i>O</i> -(4,4'-dimethoxytrityl)-2'-deoxy-β-D-ribofuranosyl)pyridine-3'- <i>O</i> -(2-cyanoethyl- <i>N,N</i> -diisopropyl)phosphoramidite	251
121	Methyl-3',5'-di- <i>O</i> -paramethoxybenzyl-2'-deoxy-D-ribose	253
122 α	2-Amino-3-methyl-5-(5'- <i>O</i> -(fluorenylmethyloxycarbonyl)-2'-deoxy-α-D-ribofuranosyl)pyridine	258
123 α	2-Amino-3-methyl-5-(5'- <i>O</i> -(4,4'-dimethoxytrityl)-2'-deoxy-α-D-ribofuranosyl)pyridine	259
123 β	2-Amino-3-methyl-5-(5'- <i>O</i> -(4,4'-dimethoxytrityl)-2'-deoxy-β-D-ribofuranosyl)pyridine	250, 275
124 α	2-[<i>N</i> -(fluorenylmethyloxycarbonyl)]amino-3-methyl-5-(3'-trimethylsilyl-5'- <i>O</i> -(4,4'-dimethoxytrityl)-2'-deoxy-α-D-ribofuranosyl)pyridine	260
124 β	2-[<i>N</i> -(fluorenylmethyloxycarbonyl)]amino-3-methyl-5-(3'-trimethylsilyl-5'- <i>O</i> -(4,4'-dimethoxytrityl)-2'-deoxy-β-D-ribofuranosyl)pyridine	264
125 α	2-[<i>N,N</i> -bis-(fluorenylmethyloxycarbonyl)]amino-3-methyl-5-(3'-trimethylsilyl-5'- <i>O</i> -(4,4'-dimethoxytrityl)-2'-deoxy-α-D-ribofuranosyl)pyridine	260
125 β	2-[<i>N,N</i> -bis-(fluorenylmethyloxycarbonyl)]amino-3-methyl-5-(3'-	264

	trimethylsilyl-5'- <i>O</i> -(4,4'-dimethoxytrityl)-2'-deoxy- β -D-ribofuranosyl)pyridine	
126 β	2-Amino-3-methyl-5-(3',5'-di- <i>O</i> -(4,4'-dimethoxytrityl)-2'-deoxy- β -D-ribofuranosyl)pyridine	262
127 β	2-[<i>N</i> -(fluorenylmethyloxycarbonyl)]amino-3-methyl-5-(5'- <i>O</i> -(4,4'-dimethoxytrityl)-2'-deoxy- β -D-ribofuranosyl)pyridine	264, 267
128 β	2-[<i>N,N</i> -bis-(fluorenylmethyloxycarbonyl)]amino-3-methyl-5-(5'- <i>O</i> -(4,4'-dimethoxytrityl)-2'-deoxy- β -D-ribofuranosyl)pyridine	264, 267, 275
129 β	2-[<i>N,N</i> -bis-(fluorenylmethyloxycarbonyl)]amino-3-methyl-5-(5'- <i>O</i> -(4,4'-dimethoxytrityl)-2'-deoxy- β -D-ribofuranosyl)pyridine-3'- <i>O</i> -(2-cyanoethyl- <i>N,N</i> -diisopropyl)phosphoramidite	269, 277
130 β	2-[<i>N,N</i> -bis-(fluorenylmethyloxycarbonyl)]amino-3-methyl-5-(3',5'-di- <i>O</i> -benzyl-2'-deoxy- β -D-ribofuranosyl)pyridine	271
131 β	2-[<i>N</i> -(fluorenylmethyloxycarbonyl)]amino-3-methyl-5-(3',5'-di- <i>O</i> -benzyl-2'-deoxy- β -D-ribofuranosyl)pyridine	271
132 β	2-[<i>N,N</i> -bis-(fluorenylmethyloxycarbonyl)]amino-3-methyl-5-(2'-deoxy- β -D-ribofuranosyl)pyridine	273
133 β	2-[<i>N</i> -(fluorenylmethyloxycarbonyl)]amino-3-methyl-5-(3'- <i>O</i> -benzyl-2'-deoxy- β -D-ribofuranosyl)pyridine	273
134 β	2-Amino-3-methyl-5-(3',5'- <i>O</i> -[1,1,3,3-tetraisopropylidisiloxane-1,3-diyl]- β -D-ribofuranosyl)pyridine	213
135	3',5'-Di- <i>O</i> -paramethoxybenzyl-2'-deoxy-D-ribose	254
136 α	3-Methyl-2-trifluoroacetamide-5-(2'-deoxy- α -D-ribofuranosyl)pyridine	257

In all cases where the same reactions have been carried out more than once, the best yields have been stated.

Methyl- 2',3',5'-tri-*O*-benzyl-D-ribose, (85)^{247,253}



A solution of hydrochloric acid in methanol (60.85 mL) [made from acetyl chloride (0.85 mL, 12 mmol) in methanol (60 mL, 1484 mmol)] was added dropwise to a solution of D-ribose (30.0 g, 200 mmol) in methanol (230 mL). The mixture was stirred at r.t for 24 hrs. Upon completion sodium bicarbonate (17.0 g, 202 mmol) was added. The mixture was filtered, concentrated *in vacuo*, co-evaporated with toluene, and then redissolved in anhydrous DMF (230 mL). Benzyl bromide (143 mL, 1200 mmol) was added and the reaction mixture cooled to 0 °C before portion wise addition of sodium hydride (60 % dispersion in mineral oil, 48.0 g, 1200 mmol). The reaction was then left to stir at r.t for 15 hrs and methanol (175 mL) was added cautiously. The mixture was reduced *in vacuo*, and then partitioned between DCM and water. The organic layers were combined and dried over anhydrous sodium sulphate, and the solvent removed *in vacuo*. Purification by column chromatography (Hex: EA 9:1, 8:1, 7:1, 0:1 v/v) afforded **85** (78.33 g, 90.1 %) as yellow oils.

Data for ***α* anomer**:

R_f (Hex: EA 7:3, A'): 0.49

LRMS: [ES⁺, MeOH] *m/z* (%): 457 ((M+Na)⁺, 37), 452 ((M+NH₄)⁺, 100).

HRMS: [ES⁺] found 457.1977 Da (M+Na⁺) calculated 457.1985 Da.

¹H NMR (300 MHz, CDCl₃): δ 7.34-7.15 (15H, m, **H^{Ar}**), 4.84 (1H, d, *J* = 1.1 Hz, **H^{1'}**), 4.59 (1H, d, *J* = 11.7 Hz, Bn-**CH**), 4.54 (1H, d, *J* = 12.1 Hz, Bn-**CH**), 4.50 (1H, d, *J* = 12.1 Hz, Bn-**CH**), 4.48 (1H, d, *J* = 12.1 Bn-**CH**), 4.46 (1H, d, *J* = 12.1 Hz, Bn-**CH**), 4.38 (1H, d, *J* = 12.1 Bn-**CH**), 4.26 (1H, ddd, *J* = 7.0, 5.5, 3.7 Hz, **H^{4'}**), 3.94 (1H, dd, *J* = 7.0, 4.7 Hz, **H^{3'}**), 3.76 (1H, dd, *J* = 4.7, 1.1 Hz, **H^{2'}**), 3.54 (1H, dd, *J* = 10.6, 4.0 Hz, **H^{5'}**), 3.44 (1H, dd, *J* = 10.6, 5.9 Hz, **H^{5'}**), 3.24 (3H, s, OCH₃).

¹³C NMR (75 MHz, CDCl₃): δ 138.5, 138.1 (**C^{Ar}**), 128.5, 128.5, 128.4, 128.1, 128.0, 127.9, 127.7, 127.6 (**CH^{Ar}**), 106.6 (**C^{1'}**), 80.7 (**C^{4'}**), 80.0 (**C^{2'}**), 78.7 (**C^{3'}**), 73.4, 72.6, 72.5 (Bn-**CH₂**), 71.5 (**C^{5'}**), 55.2 (OCH₃).

IR: ν_{max} /cm⁻¹ (neat) 3062 (w, aromatic C-H), 3029 (w, aromatic C-H), 1496 (w, aromatic C=C), 1453 (m, aromatic C=C), 1103 (m, C-O), 1065 (m, C-O), 1026 (s, C-O).

Data for **β anomer**:

R_f (Hex: EA 7:3, A⁺): 0.20

LRMS: [ES⁺, MeOH] *m/z* (%): 458 ((M(¹³C)+Na)⁺, 27), 457 ((M+Na)⁺, 100), 452 ((M+NH₄)⁺, 50).

HRMS [ES⁺] found 457.1985 Da (M+Na⁺) calculated 457.1985 Da.

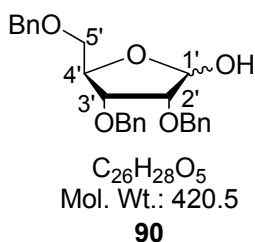
¹H NMR (300 MHz, CDCl₃): δ 7.40-7.22 (15H, m, **H^{Ar}**), 4.89 (1H, d, *J* = 4.0 Hz, **H^{1'}**), 4.71 (1H, d, *J* = 12.8 Hz, Bn-**CH**), 4.67 (1H, d, *J* = 13.1 Hz, Bn-**CH**), 4.62 (1H, d, *J* = 12.5 Hz, Bn-**CH**), 4.59 (1H, d, *J* = 12.4 Hz, Bn-**CH**), 4.51 (1H, d, *J* = 12.1 Hz, Bn-**CH**), 4.45 (1H, d, *J* = 12.1, Bn-**CH**), 4.27 (1H, dd, *J* = 6.9, 4.0 Hz, **H^{4'}**), 3.85 (1H, dd, *J* = 6.9, 3.3 Hz, **H^{3'}**), 3.80 (1H, dd, *J* = 6.6, 4.0 Hz, **H^{2'}**), 3.49 (3H, s, OCH₃), 3.45 (1H, dd, *J* = 10.6, 4.4 Hz, **H^{5'}**), 3.39 (1H, dd, *J* = 10.6, 4.4 Hz, **H^{5'}**).

¹³C NMR (75 MHz, CDCl₃): δ 138.2, 138.1, 138.0 (**C^{Ar}**), 128.5, 128.4, 128.2, 127.9, 127.8 (**CH^{Ar}**), 102.7 (**C^{1'}**), 82.2 (**C^{4'}**), 78.1 (**C^{2'}**), 75.3 (**C^{3'}**), 73.6, 72.7, 72.5 (Bn-**CH₂**), 70.4 (**C^{5'}**), 55.7 (OCH₃).

IR: $\nu_{\text{max}}/\text{cm}^{-1}$ (neat) 3062 (w, aromatic C-H), 3029 (w, aromatic C-H), 1496 (w, aromatic C=C), 1453 (m, aromatic C=C), 1104 (m, C-O), 1038 (s, C-O), 1026 (s, C-O).

Analytical data was consistent with that reported in the literature.²⁰¹

2',3',5'-Tri-*O*-benzyl-D-ribose, (90)²⁰¹



85 (78.33 g, 180.3 mmol) was stirred in acetic acid solution (80 % in water, 355 mL). Concentrated sulphuric acid (11.3 mL) was added and the mixture heated to reflux (120 °C) for 30 mins. The acetic acid was then removed *in vacuo*. DCM (355 mL) was added to the reaction mixture after it was cooled to r.t, neutralisation (pH 7-8) by sodium bicarbonate followed. The reaction mixture was then partitioned between DCM and water. The organic layers were combined and dried over anhydrous sodium sulphate. The solvent was then removed *in vacuo* to give a red-brown oil (75.8 g). The crude mixture was initially purified by filtration through a silica plug (EA: Hexane 2:3) to remove acidic and salt impurities. Further purification by column chromatography (Hex: EA 4:1, 3:1, 1:1 v/v) afforded the lactol intermediate, **90** (50.0 g, 65.9 %) as an orange-brown oil.

Data for **α anomer**:

R_f (Hex: EA 7:3, A'): 0.18

Data for **β anomer**:

R_f (Hex: EA 7:3, A'): 0.03

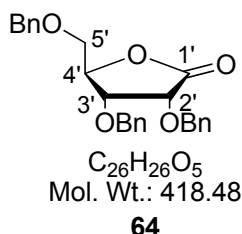
Data for **anomeric mixture**:

LRMS: [ES⁺, CH₃OH] *m/z* (%): 863 ((2M+Na)⁺, 6), 444 ((M(¹³C)+Na)⁺, 27), 443 ((M+Na)⁺, 100), 438 ((M+NH₄)⁺, 69).

HRMS: [ES⁺] found 443.1824 Da (M+Na⁺) calculated 443.1829 Da.

IR: ν_{max} /cm⁻¹ (neat); 3415 (w br, O-H), 3030 (w, aromatic C-H), 2864 (w br, O-H), 1496 (w, aromatic C=C), 1453 (m, aromatic C=C), 1240 (w, C-O), 1206 (w, C-O), 1075 (m, C-O), 1025 (m, C-O).

2',3',5'-Tri-*O*-benzyl-D-ribo-1,4-lactone, (64)²⁴⁸



90 (47.0 g, 111.7 mmol) was dissolved in anhydrous DCM (250 mL), molecular sieves were added, followed by addition of *N*-methylmorpholine-*N*-oxide (20.9 g, 178.3 mmol). The reaction mixture was then cooled to 0 °C before adding tetra-*n*-propylammonium perruthenate (VII) (2.1 g, 5.95 mmol). The reaction mixture was left to warm to r.t and stirred for 4 hrs, after which it was initially purified through a silica plug eluting with hexane. The filtrate was concentrated *in vacuo*, and upon purification by column chromatography (Hex: EA 8:1, 6:1, 4:1, 3:1 v/v) **64** was isolated as a light yellow solid (39.0 g, 83.4 %).

R_f (Hex: EA 7:3, A'): 0.41

mp: 57-59 °C, lit²⁴⁹ 54-55 °C (diisopropyl ether).

LRMS: [ES⁺, MeOH] *m/z* (%): 859 ((2M+Na)⁺, 33), 854 ((2M+NH₄)⁺, 48), 441 ((M+Na)⁺, 46), 436 ((M+NH₄)⁺, 100), 419 ((M+H)⁺, 7).

HRMS: [ES⁺] found 441.1675 Da (M+Na⁺) calculated 441.1672 Da.

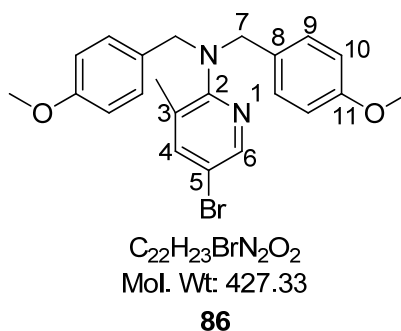
¹H NMR (300 MHz, CDCl₃): δ 7.41-7.28 (15H, m, **H^{Ar}**), 4.97 (1H, d, *J* = 12.1 Hz, Bn-CH), 4.76 (1H, d, *J* = 12.1 Hz, Bn-CH), 4.72 (1H, d, *J* = 11.0 Hz, Bn-CH), 4.58-4.54 (2H, m, **H^{4'}** plus Bn-CH), 4.52 (1H, d, *J* = 13.2 Hz, Bn-CH), 4.43 (1H, d, *J* = 11.7 Hz, Bn-CH), 4.40 (1H, d, *J* = 5.9 Hz, **H^{2'}**), 4.12 (1H, dd, *J* = 5.5, 1.8 Hz, **H^{3'}**), 3.68 (1H, dd, *J* = 11.0, 2.9 Hz, **H^{5'}**), 3.57 (1H, dd, *J* = 11.0, 2.9 Hz, **H^{5'}**).

¹³C NMR (75 MHz, CDCl₃): δ 128.7, 128.4, 128.1, 127.8 (CH^{Ar}), 81.9 (C^{4'}), 75.1 (C^{3'}), 73.9 (C^{2'}), 72.6, 72.4, 72.2 (Bn-CH₂), 68.9 (C^{5'}).

IR: ν_{max} /cm⁻¹ (neat) 3061 (w, aromatic C-H), 3032 (w, aromatic C-H), 2913 (w, alkane C-H), 2867 (w, alkane C-H), 1779 (s, C=O), 1453 (m, aromatic C=C), 1360 (m, C-O), 1101 (m, C-O), 1047 (s, C-O), 1026 (s, C-O).

Analytical data was consistent with that reported in the literature.^{201,248}

5-Bromo-2-[*N,N*-bis(4-methoxybenzyl)]amino-3-methylpyridine, (86)²⁰²



To a solution of 2-Amino-5-bromo-3-methylpyridine (20 g, 107.5 mmol) in anhydrous DMF (430 mL) was added 4-methoxybenzyl chloride (36.4 mL, 268.7 mmol) and the solution cooled to 0 °C for 30 mins before careful portion wise addition of sodium hydride (60 % dispersion in mineral oil, 10.7 g, 268.7 mmol). The reaction mixture was then left to stir at r.t for 4 hrs 35 mins. Upon completion water (5 mL) was added dropwise until the release of H₂ gas ceased. Solvent was then removed under high vacuum. The residue was redissolved in DCM and quenched with water. The organic

phases were combined, dried over anhydrous sodium sulphate and solvent removed *in vacuo*. Recrystallisation proceeded using the following method. The residue was redissolved in 95 % ethanol/water (450 mL), activated charcoal was added (4.6 g) and the mixture stirred at 105 °C for ~1 hr. The mixture was then filtered over celite (glassware had been kept in the oven for over 2 hrs). The filtrate was allowed to cool to r.t and crystallisation followed after leaving overnight at 4 °C. The crystals were filtered and washed with ice cold ethanol, and left to dry under high vacuum to give **86** (> 35 g, > 76.4 %) as white crystals.

R_f (Hex: EA 1:1, A'): 0.62

mp: 83-85 °C

LRMS: [ES⁺, CH₃CN] *m/z* (%): 451 ((M(⁸¹Br)+Na)⁺, 40), 449((M(⁷⁹Br)+Na)⁺, 38), 430 ((M(¹³C)(⁸¹Br)+H)⁺, 21), 429 ((M(⁸¹Br)+H)⁺, 99), 428 ((M(¹³C)(⁷⁹Br)+H)⁺, 22), 427 ((M(⁷⁹Br)+H)⁺, 100).

HRMS: [ES⁺] Found 427.1018 Da (M(Br⁷⁹)+H)⁺ calculated 427.1016 Da.

Found 429.1018 Da (M(Br⁸¹)+H)⁺ calculated 429.1021 Da.

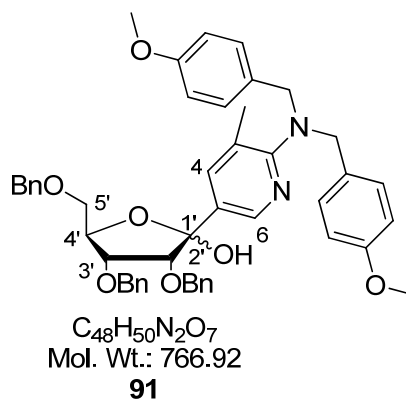
¹H NMR (300 MHz, CDCl₃): δ 8.16 (1H, d, *J* = 2.0 Hz, **H⁶**); 7.51 (1H, d, *J* = 2.0 Hz, **H⁴**); 7.14 (4H, d, *J* = 8.5 Hz, **H⁹**); 6.81 (4H, d, *J* = 8.6 Hz, **H¹⁰**); 4.21 (4H, s, **H⁷**); 3.78 (6H, s, OCH₃), 2.36 (3H, s, CH₃).

¹³C NMR (75 MHz, CDCl₃): δ 160.3 (**C²**), 158.7 (**C¹¹**), 146.0 (**C⁶**), 141.7 (**C⁴**), 130.7 (**C⁸**), 129.6 (**C⁹**), 127.8 (**C⁵**), 113.8 (**C¹⁰**), 55.3 (OCH₃), 53.8 (**C⁷**), 18.8 (CH₃).

IR: ν_{max}/cm⁻¹ (neat) 3000 (w, C-H aromatic), 2954 (w, C-H), 2929 (w, C-H), 2834 (w, C-H), 1509 (s, C=C aromatic), 1463 (m, C=C aromatic), 1448 (m, C=C aromatic), 1364 (m, C-N), 1244 (s, C-N), 1167 (s, C-N), 1033 (s, C-O).

Analytical data was consistent with that reported in the literature.^{201,202}

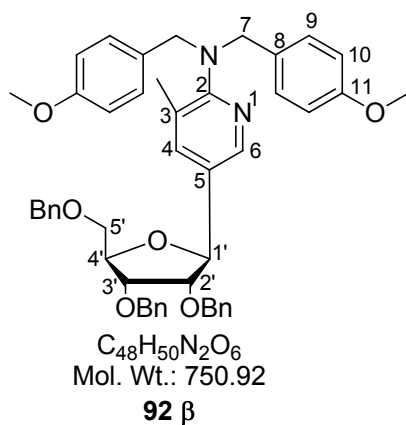
2-[*N,N*-bis(4-methoxybenzyl)]amino-3-methyl-5-(1'-hydroxyl-2',3',5'-tri-*O*-benzyl-D-ribofuranosyl)pyridine, (91**)**¹⁵⁸



86 (36.0 g, 84.2 mmol) was dissolved in anhydrous THF (180 mL) and cooled to $-78\text{ }^{\circ}\text{C}$ under an argon atmosphere before adding *n*-butyllithium (1.6 M solution in hexanes, 54.0 mL, 86.7 mmol). The mixture changed colour from yellow to orange-yellow. The reaction mixture was stirred for 1 hr 30 mins at $-78\text{ }^{\circ}\text{C}$ before adding **64** (38.0 g, 90.3 mmol) dropwise *via* cannular as a solution in THF (135 mL). The reaction mixture was stirred for a further 30 mins at $-78\text{ }^{\circ}\text{C}$ before allowing to warm to $0\text{ }^{\circ}\text{C}$, the reaction mixture was left to stir for 1 hr. Saturated sodium bicarbonate (750 mL) was added and the mixture extracted with diethyl ether. The organic layers were combined, dried over anhydrous sodium sulphate and the solvent removed *in vacuo* to give an orange-brown oil, **91**. **91** was left to dry under high vacuum overnight.

R_f (Hex: EA 7:3, B⁺): 0.20

2-[N,N-bis(4-methoxybenzyl)]amino-3-methyl-5-(2',3',5'-tri-O-benzyl-β-D-ribofuranosyl)pyridine, (92 β)¹⁵⁸

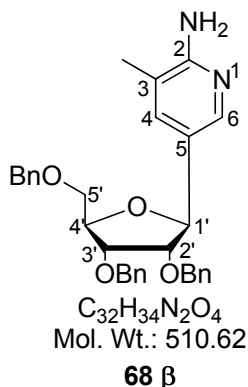


Note: Equivalents of reagents are with respect to 86.

91 was dissolved in anhydrous DCM (240 mL) and cooled to - 78 °C under an argon atmosphere before adding triethylsilane (81 mL, 504 mmol) and boron trifluoride diethyl etherate (63 mL, 504 mmol). The mixture was left to stir at - 78 °C for 4 hrs, 0 °C for 2 hrs, before leaving to warm to r.t overnight (16 hrs). Saturated sodium bicarbonate solution (120 mL) was added to the reaction mixture, and left to stir for 1 hr. The reaction mixture was then extracted with DCM. The organic layers were combined, dried over anhydrous sodium sulphate and solvent removed *in vacuo* to give a brown oil, **92 β**.

R_f (Hex: EA 7:3, B'): 0.36

2-Amino-3-methyl-5-(2',3',5'-tri-*O*-benzyl- β -D-ribofuranosyl)pyridine, (68** β)^{201,202}**



*Note: Equivalents of reagents are with respect to **86**.*

92 β was dissolved in DCM (150 mL) and to this trifluoroacetic acid (150 mL) was added, the reaction mixture was left to stir at r.t for 20 hrs. On completion the mixture was evaporated to dryness and the residue dissolved in DCM (900 mL) and washed extensively with sat. $NaHCO_3$ (aq) (1500 mL). The organic layers were combined, dried over anhydrous sodium sulphate and solvent removed *in vacuo* to give a brown oil. Purification by column chromatography (Hex: EA 1:1, 2:8, 0:1 v/v) afforded **68 β** as an orange-brown oil (25.4 g, 49.8 mmol, 59.1 %).

*Note: Final percentage yield is with respect to the starting material **86** i.e. over the 3-step synthesis.*

R_f (EA, B'): 0.21

LRMS: $[ES^+, CH_3CN]$ m/z (%): 533 ($(M+Na)^+$, 5), 512 ($(M(^{13}C)+H)^+$, 31), 511 ($(M+H)^+$, 100).

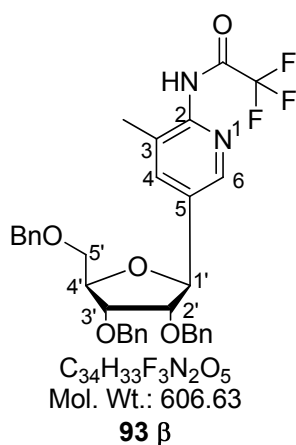
1H NMR (300 MHz, $CDCl_3$): δ 7.93 (1H, d, $J = 1.8$ Hz, H^6), 7.38-7.17 (16H, m, H^4 plus H^{Ar}), 4.89 (1H, d, $J = 7.3$ Hz, $H^{1'}$), 4.67 (1H, d, $J = 12.1$ Hz, Bn-CH), 4.62 (1H, d, $J = 12.1$ Hz, Bn-CH), 4.60 (1H, d, $J = 12.1$ Hz, Bn-CH), 4.55 (1H, d, $J = 11.7$ Hz, Bn-CH), 4.53 (1H, d, $J = 12.1$ Hz, Bn-CH), 4.48 (2H, br s, NH_2), 4.44 (1H, d, $J = 12.1$ Hz, Bn-CH), 4.31 (1H, dd, $J = 7.3, 3.7$ Hz, $H^{4'}$), 4.04 (1H, dd, $J = 5.1, 3.3$ Hz, $H^{3'}$), 3.82

(1H, dd, $J = 7.3, 4.4$ Hz, $\mathbf{H}^{2'}$), 3.66 (1H, dd, $J = 10.6, 4.0$ Hz, $\mathbf{H}^{5'}$), 3.61 (1H, dd, $J = 10.3, 3.7$ Hz, $\mathbf{H}^{5'}$), 2.18 (3H, s, \mathbf{CH}_3).

^{13}C NMR (75 MHz, CDCl_3): δ 156.9 (\mathbf{C}^2), 145.2 (\mathbf{C}^5), 143.8 (\mathbf{C}^6), 138.3, 138.1, 137.8 (\mathbf{C}^{Ar}), 136.5 (\mathbf{C}^4), 128.5, 128.4, 128.2, 128.0, 127.9, 127.8, 127.7 (\mathbf{CH}^{Ar}), 126.1 (\mathbf{C}^{Ar}), 83.3 ($\mathbf{C}^{2'}$), 82.1 ($\mathbf{C}^{4'}$), 80.5 ($\mathbf{C}^{1'}$), 77.7 ($\mathbf{C}^{3'}$) 73.6, 72.5, 72.1 (Bn- CH_2), 70.7 ($\mathbf{C}^{5'}$), 17.2 (\mathbf{CH}_3).

Analytical data was consistent with that reported in the literature.^{158,201}

3-Methyl-2-trifluoroacetamide-5-(2',3',5'-tri-*O*-benzyl- β -D-ribofuranosyl)pyridine, (93 β)²⁵⁴

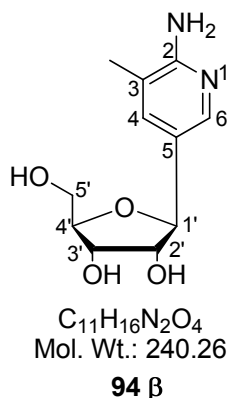


A solution of **68 β** (1.6 g, 3.13 mmol) in anhydrous pyridine (5.8 mL, 71.7 mmol), was cooled to 0 °C under an argon atmosphere, before dropwise addition of trifluoroacetic anhydride (0.5 mL, 3.45 mmol). The reaction mixture was left to warm to r.t and stirred for 17 hrs. Upon completion, the reaction mixture was quenched with saturated potassium chloride (100 mL) and extracted with DCM (300 mL). The organic layers were combined and dried over anhydrous sodium sulphate; solvent was reduced *in vacuo* to give a red-brown oil. Purification by column chromatography (Hex: EA 1:0, 3:1, 2:1, 1:1 v/v) afforded **93 β** (2.05 g, quantitative yield) as a yellow-green oil.

\mathbf{R}_f (DCM: MeOH 97:3, B β): 0.47

LRMS: [ES⁺, CH₃CN] *m/z* (%): 630 ((M(¹³C)+Na)⁺, 37), 629 ((M+Na)⁺, 100), 607 ((M+H)⁺, 35).

2-Amino-3-methyl-5-(β-D-ribofuranosyl)pyridine, (94 β)¹⁵⁸



A solution of **93 β** (2.0 g, 3.30 mmol) in anhydrous DCM (20 mL) was cooled to – 78 °C under an argon atmosphere; boron tribromide (1.0 M solution in DCM, 14.85 mL, 14.85 mmol) was added dropwise. The reaction mixture was left to stir at – 78 °C for 4 hrs. Methanol (10 mL) was added cautiously and the reaction mixture left to warm to r.t and stirred overnight. The solvent was removed *in vacuo* and the residue redissolved in methanol (2 mL). DCM (300 mL) was added portion-wise and a precipitate, **94 β** started to form. The solvent was removed via syringe and the white-grey solid, **94 β** (0.79 g, quantitative) fully formed after storage in freezer overnight.

R_f (DCM: MeOH 9:1, A⁺): 0.25

LRMS: [ES⁺, CH₃CN] *m/z* (%): 242 ((M(¹³C)+H)⁺, 16), 241 ((M+H)⁺, 100).

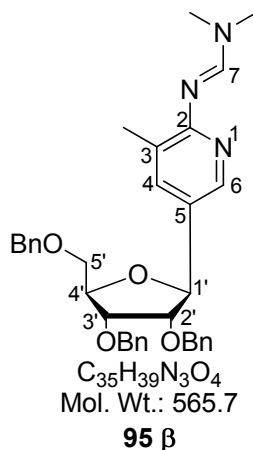
HRMS: [ES⁺] found 241.1187 Da (M+H⁺) calculated 241.1183 Da.

¹H NMR (300 MHz, CD₃OD): δ 7.89 (1H, s, **H**⁶), 7.83 (1H, s, **H**⁴), 4.62 (1H, d, *J* = 7.3 Hz, **H**^{1'}), 4.11 (1H, dd, *J* = 5.1, 3.3 Hz, **H**^{3'}), 4.00 (1H, dd, *J* = 7.7, 4.0 Hz, **H**^{4'}), 3.89 (1H, dd, *J* = 7.3, 5.5 Hz, **H**^{2'}), 3.79 (1H, dd, *J* = 12.1, 3.7, **H**^{5'}), 3.72 (1H, dd, *J* = 12.1, 4.4, **H**^{5'}), 2.27 (3H, s, CH₃).

¹³C NMR (75 MHz, CD₃OD): δ 155.2 (C²), 142.2 (C⁶), 133.1 (C⁴), 127.5 (C⁵), 123.4 (C³), 87.2 (C^{4'}), 81.6 (C^{1'}), 78.5 (C^{2'}), 72.9 (C^{3'}), 63.3 (C^{5'}), 16.8 (CH₃).

IR: ν_{max}/cm⁻¹ (neat) 3180 (s, br, O-H and N-H), 3210 (s, br, O-H and N-H), 3120 (s, br, O-H and N-H), 3075 (s, br, O-H and N-H), 3000 (w, aromatic C-H), 1668 (s, aromatic C=C), 1625 (s, aromatic C=C), 1574 (m, aromatic C=C), 1469 (w, aromatic C=C), 1454 (w, aromatic C=C), 1038 (s, C-O), 1024 (s, C-O).

2-*N,N*-[(dimethylamino)methylidene]-3-methyl-5-(2',3',5'-tri-*O*-benzyl-β-D-ribofuranosyl)pyridine, (95 β)²⁵⁵



To a solution of **68 β** (12.0 g, 23.5 mmol) in anhydrous methanol (50 mL), was added *N,N*-dimethylformamide dimethyl acetal (15.6 mL, 117.6 mmol) and the reaction mixture heated to reflux (80 °C) for 4 hrs after which, another addition of *N,N*-dimethylformamide dimethyl acetal (15.6 mL, 117.6 mmol) was added. Completion occurred after 18 hrs reflux as confirmed by MS; the solvent was removed *in vacuo*. Purification by column chromatography (DCM: MeOH 1:0, 99:1, 98:2, 97:3 v/v) afforded **95 β** (9.0 g, 67.8 %) as a yellow oil.

R_f (EA): 0.41

LRMS: [ES⁺, CH₃CN] *m/z* (%): 568 ((M(2 ¹³C)+H)⁺, 6), 567 ((M(¹³C)+H)⁺, 37), 566 ((M+H)⁺, 100).

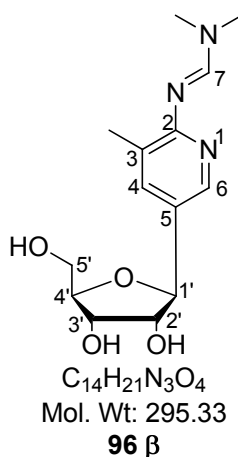
HRMS: [ES⁺] found 566.3003 Da (M+H⁺) calculated 566.3013 Da.

¹H NMR (400 MHz, CDCl₃): δ 8.27 (1H, s, **H**⁷), 8.01 (1H, d, *J* = 2.5 Hz, **H**⁶); 7.29-2.10 (16H, m, **H**⁴ plus **H**^{Ar}), 4.85 (1H, d, *J* = 7.6 Hz, **H**^{1'}), 4.57 (1H, d, *J* = 12.0 Hz, Bn-CH), 4.51 (1H, d, *J* = 12.0 Hz, Bn-CH), 4.46 (1H, d, *J* = 12.0 Hz, Bn-CH), 4.40 (1H, d, *J* = 12.1 Hz, Bn-CH), 4.36 (1H, d, *J* = 12.0 Hz, Bn-CH), 4.23 (1H, dd, *J* = 7.6, 3.5 Hz, **H**^{4'}), 3.93 (1H, dd, *J* = 5.5, 3.5 Hz, **H**^{3'}), 3.74 (1H, dd, *J* = 7.5, 5.5 Hz, **H**^{2'}), 3.57 (1H, dd, *J* = 10.5, 4.0 Hz, **H**^{5'}), 3.52 (1H, dd, *J* = 10.6, 4.0 Hz, **H**^{5'}), 3.00 (6H, s, NCH₃), 2.13 (3H, s, CH₃).

¹³C NMR (100 MHz, CDCl₃): δ 160.3 (**C**²), 154.1 (**C**⁷), 143.8 (**C**⁶), 138.0 (**C**⁴), 137.8, 137.6, 135.8 (**C**^{Ar}), 128.9, 128.3, 128.2, 128.2, 128.1, 127.9, 127.7, 127.6, 127.6, 127.5, 127.4, 127.4, 126.0 (CH^{Ar}), 83.2 (**C**^{2'}), 81.7 (**C**^{4'}), 80.5 (**C**^{1'}), 77.5 (**C**^{3'}), 73.3, 72.1, 71.8 (Bn-CH₂), 70.4 (**C**^{5'}), 40.3, 34.2 (NCH₃), 17.5 (CH₃).

Analytical data was consistent with that reported in the literature.²⁵⁵

2-*N,N*-[(dimethylamino)methylidene]-3-methyl-5-(β-D-ribofuranosyl)pyridine, (96** β)**



Method 1¹⁵⁸

A solution of **95 β** (9.09 g, 16.1 mmol) in anhydrous DCM (70 mL) was cool - 78 °C under an argon atmosphere, before the careful addition of BBr₃ (1.0 M in DCM, 72 mL,

72.0 mmol). The reaction mixture was left to stir at - 78 °C for 3 hrs before the careful addition of methanol (70 mL). The reaction mixture was left to warm to - 68 °C for 2 hrs and then r.t for 3 hrs before the removal of solvent *in vacuo*. The residue was redissolved in water (100 mL) and extracted with DCM (50 mL). The aqueous layer was isolated and co-evaporated with anhydrous pyridine. The residue was dried under high vacuum over dry KOH. A pink-brown foam **96 β** (4.7 g, quantitative) had formed.

Method 2²⁵⁵

A solution of **94 β** (0.58 g, 2.40 mmol) in anhydrous methanol (6 mL) was heated to 45 °C, *N,N*-dimethylformamide dimethyl acetal (3.2 mL, 24 mmol) was added (the reaction mixture changed from a colourless solution to a yellow solution). The reaction mixture was heated to reflux (80 °C) for 5 mins. Solvent was removed *in vacuo*. Purification by column chromatography (DCM: MeOH 97:3, 9:1, 85:15 v/v) afforded **96 β** (0.55 g, 77.1 %) as a light yellow residue.

R_f (DCM: MeOH 8:2, B⁺): 0.38

LRMS: [ES⁺, CH₃CN] *m/z* (%): 297 ((M(¹³C)+H)⁺, 17), 296 ((M+H)⁺, 100).

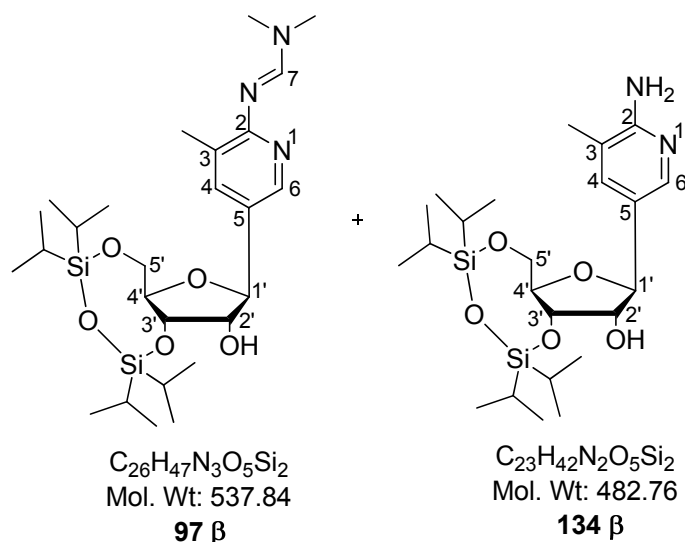
HRMS: [ES⁺] found 296.1604 Da (M+H⁺) calculated 296.1605 Da.

¹H NMR (300 MHz, DMSO): δ 8.53 (1H, s, **H**⁷), 8.09 (1H, s, **H**⁶), 7.66 (1H, d, *J* = 1.5 Hz, **H**⁴), 4.95 (3H, br s, **OH**^{2'}, **OH**^{3'} plus **OH**^{5'}), 4.54 (1H, d, *J* = 7.3 Hz, **H**^{1'}), 3.92 (1H, dd, *J* = 5.1, 3.3 Hz, **H**^{3'}), 3.81 (1H, dd, *J* = 7.7, 4.0 Hz, **H**^{4'}), 3.74 (1H, br t, *J* = 6.6 Hz, **H**^{2'}), 3.57 (1H, dd, *J* = 12.4, 4.7 Hz, **H**^{5'}), 3.51 (1H, dd, *J* = 12.1, 4.8 Hz, **H**^{5'}), 3.21 (3H, s, **NCH**₃), 3.13 (3H, s, **NCH**₃), 2.28 (3H, s, **CH**₃).

¹³C NMR (75 MHz, DMSO): δ 154.8 (**C**⁷), 141.6 (**C**⁶), 137.8 (**C**⁴), 131.7 (**C**^{Ar}), 125.5 (**C**^{Ar}), 85.4 (**C**^{4'}), 80.4 (**C**^{1'}), 77.2 (**C**^{2'}), 71.3 (**C**^{3'}), 61.9 (**C**^{5'}), 35.4 (**NCH**₃), 34.2 (**NCH**₃), 17.1 (**CH**₃).

Analytical data was consistent with that reported in the literature.^{158,201,255}

2-*N,N*-[(dimethylamino)methylidene]-3-methyl-5-(3',5'-*O*-[1,1,3,3-tetraisopropylidisiloxane-1,3-diyl]- β -D-ribofuranosyl)pyridine, (**97** β); 2-amino-3-methyl-5-(3',5'-*O*-[1,1,3,3-tetraisopropylidisiloxane-1,3-diyl]- β -D-ribofuranosyl)pyridine, (**134** β)¹⁵⁸



A solution of **96** β (4.7 g, 15.9 mmol) in anhydrous pyridine (100 mL) was cooled to 0 °C before the dropwise addition of 1,3-dichloro-1,1,3,3-tetraisopropyl disiloxane (7.6 mL, 23.9 mmol). The reaction mixture was left to stir at 0 °C for 1 hr, followed by r.t for 4 hrs. Solvent was then removed under high vacuum and the residue redissolved in DCM (100 mL) and washed with sat. NaHCO_3 (aq) (50 mL). The aqueous layer was further extracted with DCM. The organic layers were combined and dried over anhydrous sodium sulphate. Solvent was removed *in vacuo* to give a brown oil. Purification by column chromatography (Hex: EA 3:1, 2:1, 1:1, 1:2, 1:3 v/v) afforded **97** β (trans: cis 4:1, 4.86 g, 9.04 mmol) and **134** β (0.223 g, 0.46 mmol) as light yellow oils (59.7 % overall yield).

Data for isomeric mixture 97 β :

R_f (Hex: EA 1:1): 0.23

LRMS: $[\text{ES}^+, \text{CH}_3\text{CN}]$ m/z (%): 539 ($(\text{M}^{(13}\text{C})+\text{H})^+$, 38), 538 ($(\text{M}+\text{H})^+$, 100).

HRMS: [ES⁺] found 538.3121 Da (M+H⁺) calculated 538.3127 Da.

Data for major isomer (*trans*):

¹H NMR (400 MHz, CDCl₃): δ 8.32 (1H, s, **H**⁷), 8.07 (1H, d, *J* = 2.0 Hz, **H**⁶), 7.40 (1H, d, *J* = 1.5 Hz, **H**⁴), 4.72 (1H, d, *J* = 4.0 Hz, **H**^{1'}), 4.39-4.35 (1H, m, **H**^{3'}), 4.12-3.94 (4H, m, **H**^{2'}, **H**^{4'} plus **H**^{5'}), 3.07 (6H, s, NCH₃), 2.97 (1H, br s, OH^{2'}), 2.28 (3H, s, CH₃), 1.09-1.03 (28H, m, CH₃ plus CH from TiPDS).

¹³C NMR (100 MHz, CDCl₃): δ 160.6 (**C**²), 154.5 (**C**⁷), 143.9 (**C**⁶), 135.9 (**C**⁴), 129.0, 126.4 (**C**^{Ar}), 83.9 (**C**^{1'}), 82.8 (**C**^{4'}), 77.5 (**C**^{2'}), 72.2 (**C**^{3'}), 62.8 (**C**^{5'}), 40.6, 34.6 (NCH₃), 17.8, 17.6, 17.6, 17.5, 17.3, 17.2, 17.1 (CH₃ plus TiPDS-CH₃), 13.6, 13.4, 13.0, 12.8 (TiPDS-CH).

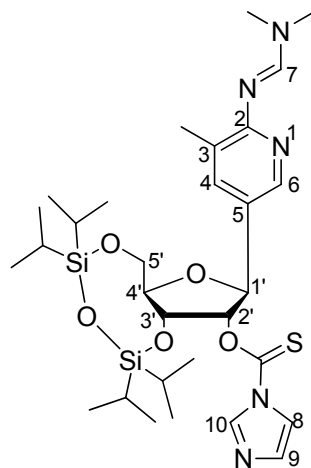
Data for minor isomer (*cis*):

¹H NMR (300MHz, CDCl₃): δ 8.33 (0.26H, s, **H**⁷), 7.95 (0.26H, s, **H**⁶), 7.30 (0.26H, s, **H**⁴), 4.66 (0.26H, d, *J* = 4.5 Hz, **H**^{1'}), 4.39-4.35 (0.26H, m, **H**^{3'}), 4.13-3.94 (1.04H, m, **H**^{2'}, **H**^{4'} plus **H**^{5'}), 3.07 (1.56H, s, NCH₃), 2.97 (0.26H, br s, OH^{2'}), 2.29 (0.78H, s, CH₃), 1.09-1.03 (7.28H, m, CH₃ plus CH from TiPDS).

¹³C NMR (100 MHz, CDCl₃): δ 160.6 (**C**²), 154.4 (**C**⁷), 144.0 (**C**⁶), 135.8 (**C**⁴), 129.0, 125.9 (**C**^{Ar}), 83.8 (**C**^{1'}), 82.8 (**C**^{4'}), 77.5 (**C**^{2'}), 72.1 (**C**^{3'}), 62.7 (**C**^{5'}), 40.6, 34.6 (NCH₃), 17.8, 17.6, 17.6, 17.5, 17.3, 17.2, 17.1 (CH₃ plus TiPDS-CH₃), 13.6, 13.4, 13.0, 12.8 (TiPDS-CH).

Analytical data was consistent with that reported in the literature.^{158,201,255}

2-*N,N*-[(dimethylamino)methylidene]-3-methyl-5-[3',5'-*O*-(1,1,3,3-tetraisopropylidisiloxane-1,3-diyl)-2'-*O*-(imidazolyl thiocarbonyl)- β -D-ribofuranosyl]pyridine, (98 β**)¹⁵⁸**



C₃₀H₄₉N₅O₅SSi₂
Mol. Wt: 647.98

98 β

To a solution of **97 β** (0.349 g, 0.65 mmol) in anhydrous DMF (3 mL) was added 1,1-thiocarbonyl diimidazole (0.3 g, 1.62 mmol) and the reaction mixture stirred at r.t under an argon atmosphere for 25 hrs. Solvent was removed under high vacuum. The residue was quenched with saturated sodium chloride solution (30 mL) and extracted with DCM (3 x 50 mL). The organic layers were combined and dried over anhydrous sodium sulphate. Solvent was removed *in vacuo* to give a yellow oil. Purification by column chromatography (DCM: MeOH 1:0, 99:1, 98:2, 97:3 v/v) afforded **98 β** (0.233 g, 55.3 %) as a light-yellow solid.

R_f (EA: NH₃: MeOH 5:1:1): 0.67

LRMS: [ES⁺, CH₃CN] *m/z* (%): 671 ((M(¹³C)+Na)⁺, 5), 670 ((M+Na)⁺, 11), 649 ((M(¹³C)+H)⁺, 36), 648 ((M+H)⁺, 100).

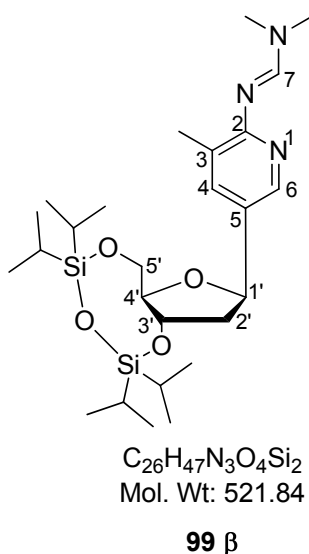
HRMS: [ES⁺] found 648.3060 Da (M+H⁺) calculated 648.3066 Da.

¹H NMR (300 MHz, CDCl₃): δ 8.32 (1H, s, **H**⁷/**H**¹⁰), 8.30 (1H, s, **H**⁷/**H**¹⁰), 8.12 (1H, d, *J* = 2.2 Hz, **H**⁶), 7.61 (1H, t, *J* = 1.5 Hz, **H**⁹), 7.46 (1H, d, *J* = 2.2 Hz, **H**⁴), 6.99 (1H, m, **H**⁸), 5.66 (1H, dd, *J* = 4.7, 1.4 Hz, **H**^{2'}), 5.10 (1H, d, *J* = 1.5 Hz, **H**^{1'}), 4.57 (1H, dd, *J* = 8.8, 5.1, **H**^{3'}), 4.15 (1H, dd, *J* = 12.8, 2.6, **H**^{5'}), 4.05-4.00 (2H, m, **H**^{4'} plus **H**^{5'}), 3.02 (6H, s, NCH₃), 2.23 (3H, s, CH₃), 1.03-0.87 (28H, m, CH₃ plus CH from TiPDS).

¹³C NMR (75 MHz, CDCl₃): δ 183.5 (C=S), 160.9 (C²), 154.6 (C⁷), 143.7 (C⁶), 136.9 (C¹⁰), 135.6 (C⁴), 131.1, 127.6 (C^{Ar}), 126.6 (C⁹), 118.2 (C⁸), 81.3 (C^{2'}), 82.2 (C^{4'}), 81.6 (C^{1'}), 69.8 (C^{3'}), 60.9 (C^{5'}), 40.6, 34.6 (NCH₃), 17.9, 17.6, 17.5, 17.4, 17.2, 17.1, 17.0 (CH₃ plus TiPDS-CH), 13.5, 13.2, 12.9, 12.8 (TiPDS-CH).

Analytical data was consistent with that reported in the literature.^{158,201,255}

2-*N,N*-[(dimethylamino)methylidene]-3-methyl-5-[3',5'-*O*-(1,1,3,3-tetraisopropylidisiloxane-1,3-diyl)-2'-deoxy-β-D-ribofuranosyl]pyridine, (99 β**)¹⁵⁸**

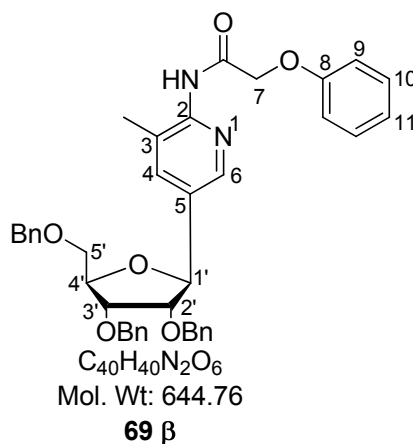


To a solution of **98 β** (0.192 g, 0.297 mmol) in anhydrous toluene (3 mL) was added azobisisobutyronitrile, AIBN (0.01 g, 0.059 mmol) and tris(trimethylsilyl)silane (0.11 mL, 0.342 mmol), under an argon atmosphere. The reaction mixture was left to stir at r.t for 1 hr while purging under argon, after which it was heated to 80 °C for 5 hrs. The reaction mixture was cooled to r.t before adding NaOH (1.0 M, 15 mL). The mixture was extracted with ethyl acetate. The organic layers were combined, dried over

sodium sulphate and the solvent removed *in vacuo*. Purification by column chromatography (Hex: EA 1:1, 0:1 v/v) caused decomposition of **99 β**. Reaction had proceeded successfully as was shown by MS monitoring throughout the reaction, producing the peak for the correct mass of the product.

LRMS: [ES⁺, CH₃CN] *m/z* (%): 1044 ((2M(¹³C)+H)⁺, 14), 1043 ((2M+H)⁺, 18), 523 ((M(¹³C)+H)⁺, 35), 522 ((M+H)⁺, 100).

3-Methyl-2-[N-(phenoxyacetyl)amino]-5-(2',3',5'-tri-*O*-benzyl-β-D-ribofuranosyl)pyridine, (69 β)¹⁵⁸



To a solution of **68 β** (4.77 g, 9.35 mmol) in anhydrous pyridine (30 mL) was added phenoxyacetic anhydride (4.55 g, 15.9 mmol) portion wise, and the reaction mixture left to stir at r.t under an argon atmosphere for 4 hrs 30 mins. Water (13 mL) was added to the reaction mixture and solvent removed under high vacuum. The residue was redissolved in ethyl acetate (150 mL) and quenched with sodium carbonate (1.0 M, 100 mL). The organic phases were combined and dried over anhydrous sodium sulphate. Solvent was removed *in vacuo*. Purification by column chromatography (DCM: EA 99:1, 98:2, 97:3, 96:4, 95:5, 9:1, 85:15, 8:2, 7:3 v/v) afforded **69 β** (4.80 g, 79.5 %) as a light-brown oil.

R_f (DCM: EA 7:3, B'): 0.57

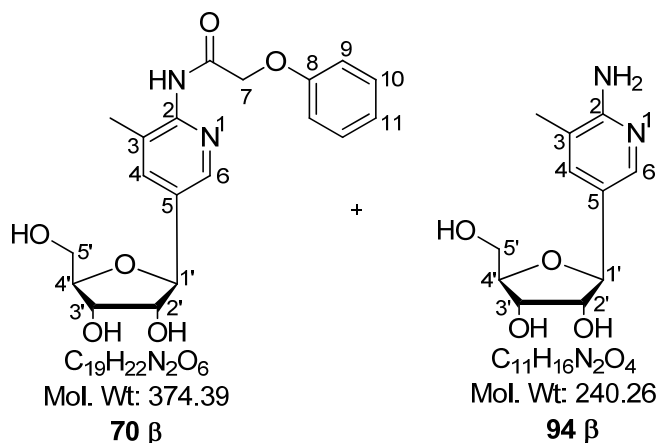
LRMS: [ES⁺, CH₃CN] *m/z* (%): 669 ((M(2 ¹³C)+Na)⁺, 11), 668 ((M(C¹³)+Na)⁺, 37), 667 ((M+Na)⁺, 100).

¹H NMR (300 MHz, CDCl₃): δ 8.58 (1H, br s, NH), 8.31 (1H, d, *J* = 2.2 Hz, **H**⁶), 7.60 (1H, d, *J* = 2.2 Hz, **H**⁴), 7.40-7.20 (17H, m, **H**^{Ar} plus **H**¹⁰), 7.07 (1H, t, *J* = 7.5 Hz, **H**¹¹); 7.02 (2H, d, *J* = 6.0 Hz, **H**⁹), 5.03 (1H, d, *J* = 7.0 Hz, **H**^{1'}), 4.70-4.43 (8H, m, Bn-CH₂ plus **H**⁷), 4.37 (1H, dd, *J* = 7.3, 3.7 Hz, **H**^{4'}), 4.06 (1H, dd, *J* = 5.1, 4.0 Hz, **H**^{3'}), 3.83 (1H, dd, *J* = 6.6, 5.1 Hz, **H**^{2'}), 3.70 (1H, dd, *J* = 10.6, 4.0, **H**^{5'}), 3.62 (1H, dd, *J* = 10.2, 3.7, **H**^{5'}), 2.16 (3H, s, CH₃).

¹³C NMR (75 MHz, CDCl₃): δ 166.7 (C=O), 157.2 (C⁸), 147.9 (C²), 144.5 (C⁶), 140.0 (C⁴), 138.1, 137.9, 137.8 (C^{Ar}), 137.5 (C⁵), 134.6 (C³), 130.0, 128.5, 128.5, 128.4, 128.2, 128.0, 128.0, 127.8, 127.7 (CH^{Ar} plus C¹⁰), 122.4 (C¹¹), 122.3 (CH^{Ar}), 114.8 (C⁹), 83.6 (C^{2'}), 82.2 (C^{4'}), 80.0 (C^{1'}), 77.4 (C^{3'}), 73.6 (Bn-CH₂), 72.6 (C^{5'}), 72.1 (Bn-CH₂), 70.4 (C⁷), 67.5 (Bn-CH₂), 18.0 (CH₃).

Analytical data was consistent with that reported in the literature.¹⁵⁸

3-Methyl-2-[N-(phenoxyacetyl)amino]-5-(β-D-ribofuranosyl)pyridine, (70 β); 2-amino-3-methyl-(β-D-ribofuranosyl)pyridine, (94 β)



A solution of **69 β** (6.65 g, 10.3 mmol) in anhydrous DCM (60 mL) was cooled to –78 °C before addition of boron tribromide (47 mL, 47 mmol) cautiously under an argon atmosphere. (On addition of BBr₃, the reaction mixture changed colour from light

brown to milky brown). The reaction mixture was left to stir at - 78 °C for 4 hrs after which methanol (23 mL) was added and the reaction mixture was left to stir overnight at r.t. Solvent was then removed *in vacuo*. The residue was quenched with DCM (60 mL) and extracted with water (120 mL). The aqueous layers were combined and neutralised with NaOH (2.0 M). Solvent was removed slightly to give an optimum concentration to allow for precipitation. Precipitation occurred after 48 hrs storage at 4 °C. The precipitate was filtered under vacuum, and then dried under high vacuum to afford **70 β** (0.837 g, 2.24 mmol) as a white solid. The filtrate was reduced *in vacuo*, and pre-adsorbed onto silica with methanol, water and acetone. The solvents were then co-evaporated with toluene. Purification by column chromatography (DCM: MeOH 9:1, 8:2, 6:4, 55:45 v/v) afforded **94 β** (1.14 g, 4.74 mmol) as a pale peach solid. The overall yield for the reaction was 67.8 %.

Data for compound **70 β**:

R_f (DCM: MeOH 9:1, A'): 0.24

LRMS: [ES⁺, MeOH] *m/z* (%): 772 ((2M+Na)⁺, 70), 398 ((M(¹³C)+Na)⁺, 22), 397 ((M+Na)⁺, 100), 376 ((M(¹³C)+H)⁺, 15), 375 ((M+H)⁺, 63).

[ES⁻, MeOH] *m/z* (%): 373 ((M+Na)⁺, 100).

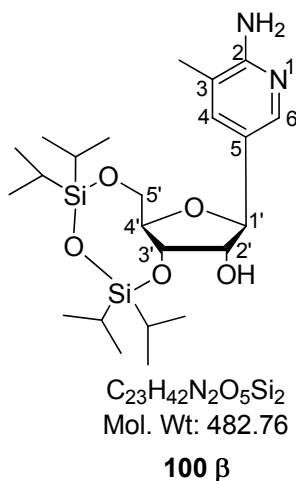
¹H NMR (400 MHz, Pyr-*d*5): δ 11.1 (1H, s, NH), 8.88 (1H, d, *J* = 1.5 Hz, **H⁶**); 8.02 (1H, d, *J* = 1.5 Hz, **H⁴**), 7.33 (2H, t, *J* = 8.5 Hz, **H¹⁰**), 7.16 (2H, d, *J* = 7.5 Hz, **H⁹**), 7.03 (1H, t, *J* = 7.0 Hz, **H¹¹**), 5.36 (1H, d, *J* = 7.5 Hz, **H^{1'}**), 5.05 (2H, s, **H⁷**), 4.84 (1H, dd, *J* = 5.5, 3.5 Hz, **H^{3'}**), 4.73 (1H, dd, *J* = 7.5, 4.0 Hz, **H^{4'}**), 4.54 (1H, dd, *J* = 7.0, 5.5 Hz, **H^{2'}**), 4.29 (1H, dd, *J* = 11.5, 4.0, **H^{5'}**), 4.23 (1H, dd, *J* = 11.6, 4.0, **H^{5'}**), 2.29 (3H, s, CH₃).

¹³C NMR (100 MHz, Pyr-*d*5): δ 168.5 (C=O), 159.0 (C⁸), 145.4 (C⁶), 138.4 (C⁴), 130.4 (C¹⁰), 122.3 (C¹¹), 115.7 (C⁹), 87.7 (C^{4'}), 82.8 (C^{1'}), 79.3 (C^{2'}), 73.4 (C^{3'}), 68.7 (C⁷), 63.7 (C^{5'}), 18.5 (CH₃).

Analytical data was consistent with that reported in the literature.¹⁵⁸

Data for compound **94** β consistent with that reported previously.

2-Amino-3-Methyl-5-[3',5'-O-(1,1,3,3-tetraisopropylidisiloxane-1,3-diyl)- β -D-ribofuranosyl]pyridine, (100** β)¹⁵⁸**



A solution of **94** β (1.14 g, 4.75 mmol) in anhydrous pyridine (16 mL) was cooled to 0 °C before dropwise addition of 1,3-dichloro-1,1,3,3-tetraisopropylidisiloxane (1.8 mL, 5.68 mmol) under an argon atmosphere. The reaction mixture was left to stir at 0 °C for 1 hr 30 mins, then r.t for 4 hrs. The solvent was co-evaporated with ethanol *in vacuo*. The residue was quenched with NaOH (1.0 M, 7 mL), and extracted with DCM (14 mL). The aqueous layer was further extracted with DCM. The organic layers were combined, dried over anhydrous sodium sulphate and the solvent removed *in vacuo* to give crude as a peach oil. Purification by column chromatography (DCM: MeOH 1:0, 99:1, 3:97 v/v) afforded **100** β (1.152 g, 50.3 %) as a peach foam.

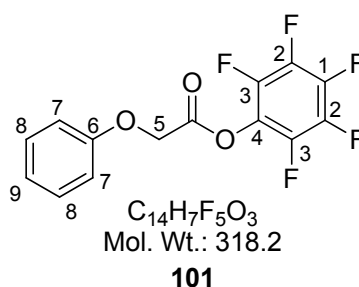
R_f (DCM: MeOH 9:1, B⁺): 0.09

LRMS: [ES⁺, MeOH] *m/z* (%): 484 ((M(¹³C)+H)⁺, 36), 483 ((M+H)⁺, 100).

¹H NMR (400 MHz, CDCl₃): δ 7.84 (1H, s, **H**⁶), 7.43 (1H, s, **H**⁴), 4.63 (1H, d, *J* = 4.5 Hz, **H**^{1'}), 4.33 (1H, t, *J* = 6.0 Hz, **H**^{3'}), 4.07 (1H, dd, *J* = 14.6, 6.0 Hz, **H**^{5'}), 4.00-3.95 (2H, m, **H**^{4'} plus **H**^{5'}), 3.87 (1H, dd, *J* = 6.0, 4.5 Hz, **H**^{2'}), 2.18 (3H, s, **CH**₃), 1.08-1.02 (28H, m, **CH**₃ and **CH** from TIPDS).

^{13}C NMR (100 MHz, CDCl_3): δ 155.7 (C^2), 138.3 (C^6), 137.7 (C^4), 125.8, 119.1 (C^{Ar}), 83.1 (C^4), 82.9 (C^1), 76.7 (C^2), 71.9 (C^3), 62.6 (C^5), 17.6 (CH_3), 17.5, 17.4, 17.2, 17.2, 17.1, 17.1, 16.9 (TIPDS- CH_3), 13.5, 13.3, 13.1, 13.0, 12.7 (TIPDS-CH).

Pentafluorophenyl 2-phenoxyethanote, (101**)**³⁰⁴



N,N'-dicyclohexylcarbodiimide, DCC (7.46 g, 36.1 mmol) was slowly added to a stirred solution of phenoxyacetic acid (5.00 g, 32.9 mmol) in DCM (50 mL) at r.t. The resulting solution was stirred for 15 mins. A solution of pentafluorophenol (6.05 g, 32.9 mmol) in DCM (50 mL) was slowly added, and the solution was stirred at r.t for a further 5 hrs. The resulting precipitate (dicyclohexylurea) was removed through filtration. The filtrate was quenched with water and extracted with DCM (3 x 100mL). The combined organic layers were dried over anhydrous sodium sulphate and solvent removed *in vacuo*. Purification by column chromatography (Hex: Et_2O 9:1 v/v) afforded **101** (7.58 g, 72.5 %) as a white solid.

R_f (Pet ether: Et_2O 9:1, A'): 0.53

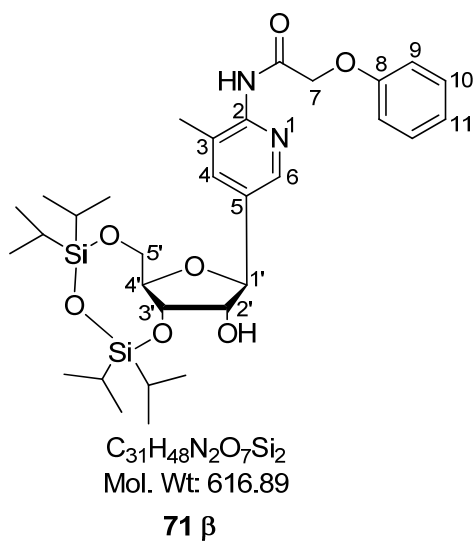
^1H NMR (400 MHz, CDCl_3): δ 7.34 (2H, t, J = 8.6 Hz, H^8), 7.06 (1H, t, J = 6.5 Hz, H^9), 6.98 (2H, d, J = 9.0 Hz, H^7), 5.00 (2H, s, H^5).

^{13}C NMR (100 MHz, CDCl_3): δ 165.4 ($\text{C}=\text{O}$), 157.1 (C^6), 141.2 (142.5 and 139.9, 2C, ddt, $^1J_{\text{C,F}}$ = 255.3 Hz, $^2J_{\text{C,F}}$ = 8.7 Hz, $^2J_{\text{C,F}}$ = 4.8 Hz, C^2 -F), 139.9 (141.2 and 138.7, 1C, dtt, $^1J_{\text{C,F}}$ = 255.8 Hz, $^2J_{\text{C,F}}$ = 13.5 Hz, $^3J_{\text{C,F}}$ = 9.7Hz, C^1 -F), 138.1 (139.3 and 136.9, 2C, dtd, $^1J_{\text{C,F}}$ = 245.4 Hz, C^3 -F), 129.9 (C^8), 124.6 (1C, m, C^4 -F), 122.6 (C^9), 114.8 (C^7), 64.8 (C^5).

¹⁹F NMR (300 MHz, CDCl₃): δ -152.5 (2F, d, ³J_{F,F} = 18.3, **F_{ortho}**), -157.2 (1F, t, ³J_{F,F} = 22.8, **F_{para}**), -162.0 (2F, t, ³J_{F,F} = 22.8, **F_{meta}**).

Analytical data was consistent with that reported in the literature.³⁰⁴

3-Methyl-2-[N-(phenoxyacetyl)amino]-5-[3',5'-O-(1,1,3,3-tetraisopropyl disiloxane-1,3-diyl)-β-D-ribofuranosyl]pyridine, (71 β)¹⁵⁸



Method 1¹⁵⁸

A solution of **70 β** (0.837 g, 2.24 mmol) in anhydrous pyridine (8 mL) was cooled to 0 °C before addition of 1,3-dichloro-1,1,3,3-tetraisopropylidisiloxane (0.84 mL, 2.69 mmol) dropwise under an argon atmosphere. The reaction mixture was left to stir at 0 °C for 1 hr, then r.t for 4 hrs. The solvent was removed *in vacuo* and the residue quenched with NaOH (1.0 M, 4 mL), and extracted with DCM (10 mL). The aqueous layer was further extracted with DCM. The organic layers were combined, dried over anhydrous sodium sulphate and the solvent removed *in vacuo* to give a mustard yellow oil. Purification by column chromatography (Hex: EA 3:1, 3:2 v/v) afforded **71 β** (0.84 g, 60.7 %) as a white foam.

Method 2²⁵⁹

To a solution of **100** **β** (1.02 g, 2.12 mmol) in anhydrous chloroform (4 mL), was added TEA (0.4 mL, 3.17 mmol). To this, a solution of pentafluorophenyl 2-phenoxyethanote, **101** (0.81 g, 2.54 mmol) in anhydrous chloroform (3 mL) was added, and the resulting reaction mixture stirred at 0 °C for 45 mins, then r.t for 21 hrs. Solvent was then removed *in vacuo*. Purification by column chromatography (Hex: EA 8:2, 7:3, 6:4 v/v) afforded **71** **β** (0.643 g, 49.2 %) as a white foam.

R_f (Hex: EA 1:1, B'): 0.29

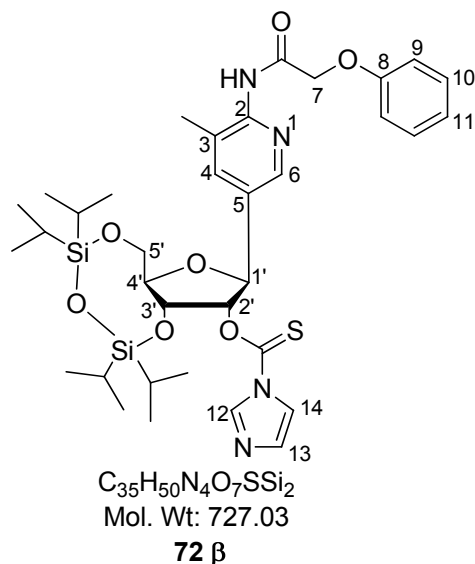
LRMS: [ES⁺, MeOH] *m/z* (%): 640 ((M(¹³C)+Na)⁺, 5), 639 ((M+Na)⁺, 48), 618 ((M(¹³C)+H)⁺, 2), 617 ((M+H)⁺, 17).

¹H NMR (400 MHz, CDCl₃): δ 8.59 (1H, br s, NH), 8.32 (1H, s, **H**⁶), 7.63 (1H, d, *J* = 1.5 Hz, **H**⁴), 7.33 (2H, dd, *J* = 9.0, 7.5 Hz, **H**¹⁰), 7.04 (1H, t, *J* = 7.5 Hz, **H**¹¹), 6.98 (2H, d, *J* = 8.0 Hz, **H**⁹), 4.82 (1H, d, *J* = 4.0 Hz, **H**^{1'}), 4.67 (2H, s, **H**⁷), 4.35 (1H, t, *J* = 6.5 Hz, **H**^{3'}), 4.13-4.00 (3H, m, **H**^{4'} plus **H**^{5'}), 3.92 (1H, m, **H**^{2'}), 3.01 (1H, br s, OH^{2'}), 2.27 (3H, s, CH₃), 1.07-0.97 (28H, m, CH plus CH₃ from TIPDS).

¹³C NMR (100 MHz, CDCl₃): δ 167.2 (C=O), 157.3 (C⁸), 148.0 (C²), 144.1 (C⁶), 137.5 (C⁴), 134.3 (C^{Ar}), 130.0 (C¹⁰), 128.1 (C^{Ar}), 122.5 (C¹¹), 114.9 (C⁹), 83.2 (C^{1'}), 82.9 (C^{4'}), 77.2 (C^{2'}), 71.7 (C^{3'}), 67.7 (C⁷), 62.4 (C^{5'}), 18.2 (CH₃), 17.6, 17.6, 17.5, 17.3, 17.2, 17.1 (TiPDS-CH₃), 13.6, 13.3, 13.2, 13.0, 12.8 (TiPDS-CH).

Analytical data was consistent with that reported in the literature.¹⁵⁸

3-Methyl-2-[N-(phenoxyacetyl)amino]-5-[3',5'-O-(1,1,3,3-tetraisopropyl disiloxane-1,3-diyl)-2'-O-(imidazolyl thiocarbonyl)-β-D-ribofuranosyl]pyridine, (72 β)¹⁵⁸



To a solution of **71 β** (0.840 g, 1.36 mmol) in anhydrous DMF (8 mL) was added 1,1-thiocarbonyl diimidazole (0.616 g, 3.41 mmol) and the reaction mixture was left to stir at r.t under an argon atmosphere for 29 hrs. Solvent was removed *in vacuo*. The reaction mixture was quenched with water (15 mL) and extracted with ethyl acetate (25 mL); the aqueous layers were further extracted with ethyl acetate (3 x 25 mL). The combined organic layers were dried over anhydrous sodium sulphate. Solvent was removed *in vacuo* to give crude as a yellow oil. Purification by column chromatography (Hex: EA 3:1, 2:1, 1:1 v/v) afforded **72 β** (0.551 g, 55.1 %) as a yellow foam.

R_f (Hex: EA 1:1, B'): 0.24

LRMS: [ES⁺, MeOH] *m/z* (%): 751 ((M(¹³C) +Na)⁺, 45), 750 ((M+Na)⁺, 95), 729 ((M(¹³C)+H)⁺, 22), 728 ((M+H)⁺, 100).

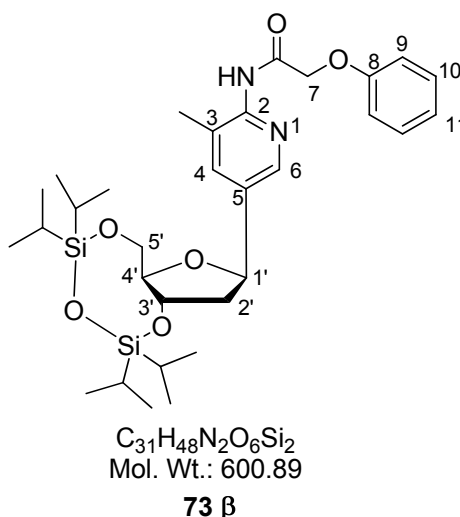
¹H NMR (400 MHz, CDCl₃): δ 8.58 (1H, br s, NH), 8.50 (1H, d, *J* = 2.0, H⁶), 8.39 (1H, s, H¹²), 7.79 (1H, d, *J* = 1.5 Hz, H⁴), 7.68 (1H, s, H¹³), 7.35 (2H, t, *J* = 8.5 Hz, H¹⁰), 7.07 (1H, d, *J* = 1.0 Hz, H¹⁴), 7.04 (1H, d, *J* = 7.5 Hz, H¹¹), 6.99 (2H, d, *J* = 8.5 Hz, H⁹), 5.73 (1H, d, *J* = 5.0 Hz, H^{2'}), 5.26 (1H, s, H^{1'}), 4.70 (2H, s, H⁷), 4.61 (1H, dd, *J* = 9.0,

5.0 Hz, $\mathbf{H}^{3'}$), 4.27-4.24 (1H, m, $\mathbf{H}^{5'}$), 4.12-4.08 (2H, m, $\mathbf{H}^{4'}$ plus $\mathbf{H}^{5'}$), 2.30 (3H, s, \mathbf{CH}_3), 1.11 –0.99 (28H, m, \mathbf{CH} plus \mathbf{CH}_3 from TiPDS).

^{13}C NMR (100 MHz, CDCl_3): δ 183.5 ($\text{C}=\text{S}$), 166.9 ($\text{C}=\text{O}$), 157.2 (C^8), 148.5 (C^2), 144.5 (C^6), 137.6 (C^4), 136.9 (C^{12}), 132.8 (C^5), 131.2 (C^{14}), 130.0 (C^{10}), 122.5 (C^{11}), 118.2 (C^{13}), 114.9 (C^9), 87.0 ($\text{C}^{2'}$), 82.3 ($\text{C}^{4'}$), 81.1 ($\text{C}^{1'}$), 69.2 ($\text{C}^{3'}$), 67.7 (C^7), 60.4 ($\text{C}^{5'}$), 18.2 (CH_3), 17.6, 17.5, 17.4, 17.2, 17.1, 17.0 (TiPDS- CH_3), 13.5, 13.1, 12.9, 12.8 (TiPDS- CH).

Analytical data was consistent with that reported in the literature.^{158,201}

3-Methyl-2-[*N*-(phenoxyacetyl)amino]-5-[3',5'-*O*-(1,1,3,3-tetraisopropylidisiloxane-1,3-diyl)-2'-deoxy- β -D-ribofuranosyl]pyridine, (73 β**)¹⁵⁸**



To a solution of **72 β** (0.306 g, 0.421 mmol) in anhydrous toluene (6.0 mL) was added azobisisobutyronitrile, AIBN (0.14 g, 0.844 mmol) followed by tris(trimethylsilyl)silane (0.15 mL, 0.485 mmol). The reaction mixture was purged under argon and stirred at r.t for 1 hr, after which it was heated to 80 °C for 5 hrs 30 mins. The reaction mixture was left to cool to r.t and solvent was removed *in vacuo*. The crude was then quenched with NaOH (1.0 M, 6 mL), and extracted with ethyl acetate (12 mL x 3). The organic layers were combined and dried over anhydrous sodium sulphate. Solvent was removed *in vacuo* to give following purification by column chromatography (Hex: EA 1:1 v/v) **73 β** (0.15 g, 61.4 %) as a yellow oil.

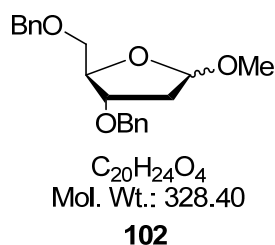
R_f (Hex: EA 1:1, A'): 0.34

LRMS: [ES⁺, MeOH] *m/z* (%): 602 ((M(¹³C)+H)⁺, 36), 601 ((M+H)⁺, 100).

¹H NMR (400 MHz, CDCl₃): δ 8.53 (1H, br s, NH), 8.24 (1H, d, *J* = 2.0 Hz, H⁶), 7.57 (1H, d, *J* = 2.0 Hz, H⁴), 7.36-7.32 (2H, m, H¹⁰), 7.05 (1H, t, *J* = 7.5 Hz, H¹¹), 6.99 (2H, dd, *J* = 8.5, 1.0 Hz, H⁹), 5.10 (1H, t, *J* = 7.5 Hz, H^{1'}), 4.68 (2H, s, H⁷), 4.53 (1H, dt, *J* = 7.5, 4.5 Hz, H^{3'}), 4.12 (1H, dd, *J* = 10.5, 2.0 Hz, H^{5'}), 3.93-3.85 (2H, m, H^{4'} plus H^{5'}), 2.39 (1H, ddd, *J* = 12.5, 7.0, 4.5 Hz, H^{2'}), 2.27 (3H, s, CH₃), 2.04 (1H, dt, *J* = 13.0, 7.5 Hz, H^{2'}), 1.09-1.03 (28H, m, CH plus CH₃ from TiPDS).

¹³C NMR (100 MHz, CDCl₃): δ 166.8 (C=O), 157.2 (C²), 147.8 (C^{Ar}), 144.2 (C⁶), 137.5 (C⁴), 136.3 (C^{Ar}), 130.0 (C¹⁰), 128.5 (C^{Ar}), 122.5 (C¹¹), 114.9 (C⁹), 86.7 (C^{4'}), 76.5 (C^{1'}), 73.2 (C^{3'}), 67.7 (C⁷), 63.5 (C^{5'}), 43.0 (C^{2'}), 18.1 (CH₃), 17.7, 17.6, 17.5, 17.4, 17.2, 17.2, 17.1 (TiPDS-CH₃), 13.7, 13.5, 13.1, 12.7 (TiPDS-CH).

Methyl-3',5'-di-*O*-benzyl-2'-deoxy-D-ribose, (102)^{247,253}



To a solution of 2'-deoxy-D-ribose (15.0 g, 111.8 mmol) in anhydrous methanol (560 mL), a solution of acetyl chloride (0.5 mL, 6.71 mmol) in anhydrous methanol (33.6 mL, 829.8 mmol) was added dropwise and the reaction mixture left to stir at r.t for 1 hr. Upon completion, NaHCO₃ (13.0 g, 155 mmol) was added until pH 7.0. The reaction mixture was then filtered. Solvent was removed *in vacuo*. The residue was co-evaporated with toluene. The residue was then redissolved in anhydrous DMF (560 mL), benzyl bromide (53.2 mL, 447.3 mmol) was added and the reaction mixture cooled to 0 °C before portion wise addition of NaH (60 % in mineral oil, 18.0 g, 447.3 mmol). The reaction mixture was then left to warm to r.t and left for 14 hrs. On

completion, the reaction mixture was cooled to 0 °C over an ice bath before careful addition of anhydrous methanol (400 mL) and the mixture evaporated to dryness followed by co-evaporation with toluene. The residue was then quenched with water and extracted with DCM (500 mL). The organic phases were combined and dried over anhydrous sodium sulphate. Solvent was removed *in vacuo*. Column chromatography (DCM: EA 1:0, 99:1, 95:5, 9:1, 8:2, 7:3 and 6:4 v/v) afforded **102** (> 34.84 g, > 95 %) as a faint yellow oil.

Data for α anomer:

R_f (Hex: EA 7:3, A'): 0.37

LRMS: [ES⁺, MeOH] *m/z* (%): 680 ((2M(¹³C)+Na)⁺, 9), 679 ((2M+Na)⁺, 33), 352 ((M(¹³C)+Na)⁺, 25), 351 ((M+Na)⁺, 100).

¹H NMR (400 MHz, CDCl₃): δ 7.33-7.28 (10H, m, **H^{Ar}**), 5.07 (1H, dd, *J* = 5.5, 1.0, **H^{1'}**), 4.58-4.47 (4H, m, Bn-CH₂), 4.24 (1H, dd, *J* = 8.5, 4.0, **H^{4'}**), 3.97 (1H, ddd, *J* = 8.0, 4.5, 3.0, **H^{3'}**), 3.54 (1H, dd, *J* = 10.0, 4.0, **H^{5'}**), 3.50 (1H, dd, *J* = 10.5, 4.5, **H^{5'}**), 3.40 (3H, s, OCH₃), 2.22 (1H, ddd, *J* = 14.1, 8.0, 5.5, **H^{2'}**), 2.02 (1H, ddd, *J* = 14.0, 2.5, 1.0, **H^{2'}**).

¹³C NMR (100 MHz, CDCl₃): δ 138.3, 138.3 (**C^{Ar}**), 128.5, 128.0, 127.8, 127.8 (**CH^{Ar}**), 105.3 (**C^{1'}**), 82.3 (**C^{4'}**), 78.7 (**C^{3'}**), 73.6, 71.7 (Bn-CH₂), 70.3 (**C^{5'}**), 55.3 (OCH₃), 39.1 (**C^{2'}**).

Data for β anomer:

R_f (Hex: EA 7:3, A'): 0.46

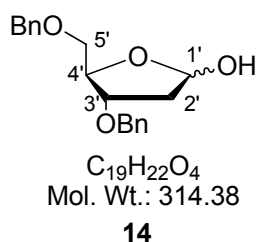
LRMS: [ES⁺, MeOH] *m/z* (%): 680 ((2M(¹³C)+Na)⁺, 9), 679 ((2M+Na)⁺, 33), 352 ((M(¹³C)+Na)⁺, 25), 351 ((M+Na)⁺, 100).

¹H NMR (400 MHz, CDCl₃): δ 7.42-7.34 (10H, m, **H^{Ar}**), 5.17 (1H, dd, *J* = 5.5, 2.5, **H^{1'}**), 4.68-4.54 (4H, m, Bn-CH₂), 4.35 (1H, td, *J* = 6.0, 3.5, **H^{4'}**), 4.22 (1H, td, *J* = 6.5, 3.5, **H^{3'}**), 3.63 (1H, dd, *J* = 9.5, 6.0, **H^{5'}**), 3.57 (1H, dd, *J* = 9.6, 6.0, **H^{5'}**), 3.36 (3H, s, OCH₃), 2.30 (1H, ddd, *J* = 13.1, 6.6, 2.0, **H^{2'}**), 2.20 (1H, dt, *J* = 13.6, 5.5, **H^{2'}**).

^{13}C NMR (100 MHz, CDCl_3): δ 138.3, 138.1 (C^{Ar}), 128.9, 128.5, 128.5, 128.3, 127.8, 127.8, 127.7 (CH^{Ar}), 105.6 ($\text{C}^{1'}$), 83.0 ($\text{C}^{4'}$), 80.1 ($\text{C}^{3'}$), 73.4 (Bn-CH_2), 72.1 ($\text{C}^{5'}$), 71.7 (Bn-CH_2), 55.0 (OCH_3), 39.5 ($\text{C}^{2'}$).

Analytical data consistent with that reported in the literature.²⁵³

3',5'-Di-*O*-benzyl-2'-deoxy-D-ribo-1,4-lactol, (14)^{261,262}



A solution of **102** (17.82 g, 54.3 mmol) in acetic acid solution (80 % acetic acid in water, 232 mL) was heated to 110 °C for 2 hrs after which the solution changed colour from faint yellow to mustard yellow. On completion, solvent was removed *in vacuo* before neutralization by sat. NaHCO_3 (aq). The reaction mixture was then extracted with DCM (3 x 500 mL). The organic phases were combined and dried over anhydrous sodium sulphate. Solvent was removed *in vacuo* to give the crude as a mustard-yellow oil (15.2 g). Column chromatography (DCM: EA 1:0, 99.5:0.5, 99:1, 98:2, 97:3, 96:4, 95:5, 94:6, 93:7, 92:8, 91:9, 9:1, 88:15, 7:3 v/v) afforded **14** (10.25 g, 60 %) as a pale yellow oil.

Data for α anomer:

R_f (Pet ether: EA 7:3, A'): 0.27

Data for β anomer:

R_f (Pet ether: EA 7:3, A'): 0.43

Data for the anomeric mixture:

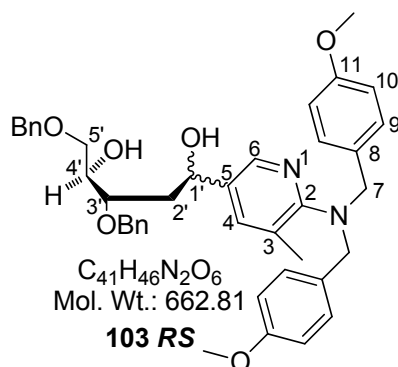
LRMS: $[\text{ES}^+, \text{MeOH}]$ m/z (%): 652 ($((2\text{M}(^{13}\text{C})+\text{Na})^+, 2)$, 651 ($((\text{M}+\text{Na})^+, 9)$, 338 ($((\text{M}(^{13}\text{C})+\text{Na})^+, 19)$, 337 ($((\text{M}+\text{Na})^+, 100)$).

HRMS: [ES⁺] Found 337.1415 Da (M+Na)⁺, calculated 337.1410 Da.

¹H NMR (400 MHz, CDCl₃): δ 7.37-7.32 (20H, m, **H^{Ar}**), 5.58 (1H, br t, *J* = 3.5 Hz, **H^{1'α}**), 5.50 (1H, d, *J* = 9.5 Hz, **H^{1'β}**), 4.62-4.46 (9H, m, Bn-CH₂ plus **H^{4'β}**), 4.29-4.26 (2H, m, **H^{3'α}** plus **H^{4'α}**), 4.13-4.11 (1H, m, **H^{3'β}**), 4.01 (1H, d, *J* = 8.0 Hz, **OHα**), 3.80 (1H, d, *J* = 9.5 Hz, **OHβ**), 3.65 (1H, dd, *J* = 10.0, 3.5 Hz, **H^{5'α}**), 3.57 (1H, dd, *J* = 10.0, 3.5 Hz, **H^{5'α}**), 3.53 (1H, dd, *J* = 10.0, 4.5 Hz, **H^{5'β}**), 3.40 (1H, dd, *J* = 10.0, 5.5 Hz, **H^{5'β}**), 2.22 (2H, m, **H^{2'α}**), 2.14 (2H, t, *J* = 3.0, **H^{2'β}**).

¹³C NMR (100 MHz, CDCl₃): δ 138.0, 137.9, 137.5, 137.4 (**C^{Arα,β}**), 128.6, 128.5, 128.5, 128.4, 128.0, 127.9, 127.9, 127.8, 127.8, 127.8, 127.7, 127.6 (**CH^{Arα,β}**), 99.3 (**C^{1'α,β}**), 83.1 (**C^{4'α}**), 82.4 (**C^{4'β}**), 80.0 (**C^{3'β}**), 79.8 (**C^{3'α}**), 73.6, 73.5, 71.5, 71.4 (Bn-CH₂), 71.3 (**C^{5'α}**), 70.4 (**C^{5'β}**), 41.5 (**C^{2'α}**), 39.2 (**C^{2'β}**).

2-[*N,N*-bis(4-methoxybenzyl)]amino-3-methyl-5-(1'-hydroxyl-2'-deoxy-3',5'-di-*O*-benzyl-D-ribofuranosyl)pyridine, (103 *RS*)²²⁰



Coupling reaction Method 1²²⁰

A solution of **86** (1.0 g, 2.35 mmol) in anhydrous THF (6.5 mL) was cooled to - 50 °C before dropwise addition of butyllithium (1.6 M solution in hexanes, 1.5 mL, 2.35 mmol). The reaction mixture was left to stir for 2 hrs 30 mins, before dropwise addition of a solution of **14** (0.3 g, 0.94 mmol) in anhydrous THF. The reaction mixture was then left to warm to r.t, after 2 hrs 30 mins reaction had not completed hence another addition of **14** (0.27 g, 0.86 mmol) in anhydrous THF to the reaction mixture. After overnight stirring the reaction was still incomplete. The reaction was stopped

however, and solvent removed *in vacuo*. The brown-orange residue was quenched with water and extracted with ethyl acetate. The organic phases were combined and dried over anhydrous sodium sulphate. Solvent was removed *in vacuo* and column chromatography (Hex: EA 8:2, 7:3, 6:4, 1:1 v/v) afforded **103 R** and **103 S** (0.11 g, 18 %) as orange-brown oils.

Coupling reaction Method 2²²⁰

A solution of **86** (53.8 g, 126.1 mmol) in anhydrous THF (180 mL) was cooled to - 35 °C before dropwise addition of n-butyllithium (1.6 M solution in hexanes, 79.0 mL, 126.1 mmol), the reaction mixture changed colour from yellow to dark brown. The reaction mixture was left to stir at - 35 °C for 30 mins, before dropwise addition of a solution of **14** (15.84 g, 50.4 mmol) in anhydrous THF (50 mL), the reaction mixture became a light red-brown colour. The reaction mixture was then left to stir at - 30 °C for 1 hr, after which it was allowed to warm to r.t for 4 hrs 30 mins. Solvent was removed *in vacuo*. The brown-orange residue was quenched with water and extracted with EA. The organic phases were combined and dried over anhydrous sodium sulphate. Solvent was removed *in vacuo* and column chromatography (DCM: EA 1:0, 95:5, 93:7, 9:1, 1:1, EA:MeOH 8:2, 7:3 v/v) afforded the epimeric mixture, **103 RS** (30.74 g, 92.0 %) as a brown foam.

Data for the epimeric mixture:

R_f (DCM: EA 8:2, A'): 0.55

HRMS: [ES⁺] Found 663.3421 Da (M+H)⁺, calculated 663.3429 Da.

Data for less polar epimer:

R_f (Hex: EA 1:1, A'): 0.35

LRMS: [ES⁺, MeOH] *m/z* (%): 685 ((M+Na)⁺, 12), 664 ((M(¹³C) +H)⁺, 40), 663 ((M+H)⁺, 100)

[ES⁻, MeOH] *m/z* (%): 663 ((M(¹³C)-H)⁻, 15), 662 ((M-H)⁻, 100).

¹H NMR (400 MHz, CDCl₃): δ 8.11 (1H, d, *J* = 2.0 Hz, **H**⁶), 7.43 (1H, d, *J* = 2.0 Hz, **H**⁴), 7.37-7.29 (10H, m, **H**^{Ar}), 7.17 (4H, d, *J* = 8.5 Hz, **H**⁹), 6.80 (4H, d, *J* = 8.5 Hz, **H**¹⁰), 4.88 (1H, br t, *J* = 6.5 Hz, **H**^{1'}), 4.63-4.59 (2H, m, Bn-CH₂), 4.56 (2H, s, BnCH₂), 4.23 (4H, s, **H**⁷), 4.03-3.98 (1H, m, **H**^{4'}), 3.81 (1H, dd, *J* = 11.0, 5.5 Hz, **H**^{3'}), 3.76 (6H, s, OCH₃), 3.65 (1H, dd, *J* = 9.5, 3.5 Hz, **H**^{5'}), 3.60 (1H, dd, *J* = 9.5, 6.5 Hz, **H**^{5'}), 3.17 (1H, br s, OH^{1'}), 2.86 (1H, br s, OH^{4'}), 2.37 (3H, s, CH₃), 2.00 (2H, br t, *J* = 6.0 Hz, **H**^{2'}).

¹³C NMR (100 MHz, CDCl₃): δ 160.9 (**C**²), 158.5 (**C**¹¹), 142.9 (**C**⁶), 138.0, 137.9 (**C**^{Ar}), 137.3 (**C**⁴), 134.2, 131.2, 130.9 (**C**^{Ar}), 129.6 (**C**⁹), 128.6, 128.6, 128.4, 128.2, 128.1, 128.0 (CH^{Ar}), 125.8 (**C**^{Ar}), 113.7 (**C**¹⁰), 77.4 (**C**^{3'}), 73.6, 72.6 (Bn-CH₂), 71.8 (**C**^{4'}), 71.1 (**C**^{5'}), 68.9 (**C**^{1'}), 55.2 (OCH₃), 53.8 (**C**⁷), 39.5 (**C**^{2'}), 18.8 (CH₃).

Data for more polar epimer:

R_f (Hex: EA 1:1, A⁺): 0.25

LRMS: [ES⁺, MeOH] *m/z* (%): 685 ((M+Na)⁺, 12), 664 ((M(¹³C) +H)⁺, 40), 663 ((M+H)⁺, 100)

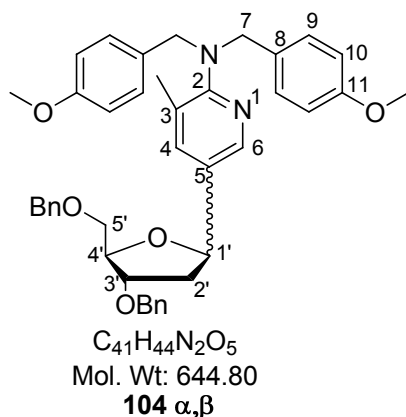
[ES⁻, MeOH] *m/z* (%): 663 ((M(¹³C)-H)⁻, 15), 662 ((M-H)⁻, 100).

¹H NMR (400 MHz, CDCl₃): δ 8.12 (1H, d, *J* = 2.5, **H**⁶), 7.43 (1H, d, *J* = 2.0, **H**⁴), 7.36-7.28 (10H, m, **H**^{Ar}), 7.17 (4H, d, *J* = 8.5, **H**⁹), 6.81 (4H, d, *J* = 8.6, **H**¹⁰), 4.89 (1H, dd, *J* = 8.5, 3.0, **H**^{1'}), 4.61-4.51 (4H, m, Bn-CH₂), 4.23 (4H, s, **H**⁷), 4.05 (1H, dd, *J* = 10.6, 5.5, **H**^{4'}), 3.76 (7H, br s, OCH₃ plus **H**^{3'}), 3.61-3.58 (2H, m, **H**^{5'}), 2.38 (3H, s, CH₃), 2.13 (1H, br dt, *J* = 15.1, 8.5, **H**^{2'}), 1.92 (1H, dt, *J* = 15.0, 3.5, **H**^{2'}).

¹³C NMR (100 MHz, CDCl₃): δ 161.0 (**C**²), 158.5 (**C**¹¹), 143.1 (**C**⁶), 137.8, 137.8 (**C**^{Ar}), 137.8 (**C**⁴), 134.2, 134.0, 131.1, 130.9 (**C**^{Ar}), 129.6 (**C**⁹), 128.6, 128.5, 128.2, 128.1, 128.0, 127.9 (CH^{Ar}), 125.8 (**C**^{Ar}), 113.6 (**C**¹⁰), 78.7 (**C**^{3'}), 73.5, 72.1 (Bn-CH₂), 71.6 (**C**^{4'}), 70.9 (**C**^{5'}), 69.5 (**C**^{1'}), 55.2 (OCH₃), 53.8 (**C**⁷), 39.0 (**C**^{2'}), 18.8 (CH₃).

Analytical data consistent with that reported in the literature.^{158,220}

2-[*N,N*-bis(4-methoxybenzyl)]amino-3-methyl-5-(3',5'-di-*O*-benzyl-2'-deoxy-D-ribofuranosyl)pyridine, (104** α,β)**²²⁰



Mitsunobu Cyclisation Method 1

A solution of **103 RS** (0.173 g, 0.261 mmol) in anhydrous benzene (5 mL) was cooled to 0 °C, Bu_3P (0.08 mL, 0.34 mmol) was added followed by TMAD (0.06 g, 0.34 mmol). The reaction mixture was left to stir at r.t for 22 hrs, after which the insoluble material was filtered and the filtrate concentrated *in vacuo*. The resulting crude oil was quenched with water and extracted with ethyl acetate. The organic layers were combined and dried over anhydrous sodium sulphate, solvent was removed *in vacuo*. Column chromatography (Hex: EA 1:0, 95:5, 9:1, 8:2 v/v) afforded **104** (0.075 g, 44.6 %) as 2 anomers ($\alpha:\beta$ 1:2) as light yellow oils.

Mitsunobu Cyclisation Method 2

A solution of **103 RS** (0.180 g, 0.27 mmol) in anhydrous benzene (1 mL) was cooled to 0 °C, Bu_3P (0.10 mL, 0.41 mmol) was added followed by DHTD (0.07 g, 0.41 mmol). The reaction mixture was left to stir at r.t for 25 hrs, after which the insoluble material was filtered and the filtrate concentrated *in vacuo*. The resulting crude oil was quenched with water and extracted with ethyl acetate. The organic layers were combined and dried over anhydrous sodium sulphate, solvent was removed *in vacuo*. Column

chromatography (Hex: EA 1:0, 95:5, 9:1 v/v) afforded **104 α** and **104 β** (α : β 2:3, 0.0945 g, 54.3 %) as yellow oils.

Mitsunobu Cyclisation Method 3

A solution of **103 *RS*** (0.215 g, 0.32 mmol) in anhydrous THF (2.0 mL) was cooled to 0 °C, Bu₃P (0.13 mL, 0.52 mmol) was added followed by DIAD (0.1 mL, 0.52 mmol). The reaction mixture was left to stir at r.t for 15 hrs 30 mins, after which solvent was removed *in vacuo*. The resulting crude oil in methanol was loaded onto a pyridinium activated dowex column (the dowex had been washed with 10 % pyridine in MeOH: H₂O 1:1). First the reduced DIAD by-product was removed in methanol (300 mL), followed by water (300 mL). The anomeric product **104 α,β** was released from the resin using 10 % NH₃ in MeOH: H₂O 1:1 (3.0 L). The pyridinium form dowex was used to avoid removal of paramethoxy benzyl groups by the acidic dowex, however the compound is not basic enough to exchange with the pyridinium ions, hence there was no separation of the reduced DIAD from the anomeric product.

Mitsunobu Cyclisation Method 4

A solution of **103 *RS*** (3.1832 g, 4.8 mmol) in anhydrous THF (24.0 mL) was cooled to 0 °C, Bu₃P (2.0 mL, 7.68 mmol) was added followed by DIAD (1.5 mL, 7.68 mmol). The reaction mixture was left to stir at r.t for 5 hrs 30 mins, after which solvent was removed *in vacuo*. Column chromatography (Hex: EA 1:0, 95:5, 9:1 v/v) afforded **104** as 2 anomers **α** (1.21 g) and **β** (2.04 g) as yellow oils; both containing the reduced DIAD by-product.

Mitsunobu Cyclisation Method 5

A solution of **103 *RS*** (12.49 g, 18.84 mmol) in anhydrous THF (86.0 mL) was cooled to 0 °C, before addition of Bu₃P (7.44 mL, 30.15 mmol) followed by TMAD (5.19 g, 30.15 mmol). The reaction mixture was left to stir at r.t for 6 hrs 15 mins, after which the reaction mixture was filtered and the filtrate concentrated *in vacuo*. The residue was quenched with water and extracted with ethyl acetate. The organic phases were

combined and dried over anhydrous sodium sulphate. Solvent was removed *in vacuo* to afford the crude (23.11 g) as a brown oil. Column chromatography (Pet ether: EA 5:1, 4:1, 3:1 v/v) afforded **104 α** and **104 β** (α:β 2:3, 10.77 g, 88.2 %) as yellow oils.

Data for the anomeric mixture:

HRMS: [ES⁺] Found 645.3324 Da (M+H)⁺, calculated 645.3323 Da.

Data for α anomer:

R_f (Hex: EA 1:1, A⁺): 0.58

¹H NMR (400 MHz, CDCl₃): δ 8.09 (1H, d, *J* = 2.0 Hz, **H⁶**), 7.59 (1H, s, **H⁴**), 7.36-7.28 (10H, m, **H^{Ar}**), 7.15 (4H, d, *J* = 8.5 Hz, **H⁹**), 6.80 (4H, d, *J* = 8.5 Hz, **H¹⁰**), 5.02 (1H, t, *J* = 8.0 Hz, **H^{1'}**), 4.63-4.52 (4H, m, Bn-CH₂), 4.36 (1H, dd, *J* = 9.0, 4.5 Hz, **H^{4'}**), 4.26 (1H, dd, *J* = 10.6, 6.5 Hz, **H^{3'}**), 4.22 (4H, s, **H⁷**), 3.78 (6H, s, OCH₃), 3.61 (2H, m, **H^{5'}**), 2.61 (1H, dt, *J* = 13.0, 6.5 Hz, **H^{2'β}**), 2.36 (3H, s, CH₃), 2.05 (1H, ddd, *J* = 13.6, 8.0, 6.0 Hz, **H^{2'α}**).

¹³C NMR (100 MHz, CDCl₃): δ 161.2 (**C²**), 158.6 (**C¹¹**), 143.5 (**C⁶**), 138.3 (**C⁴**), 138.1, 137.9, 131.8, 131.1 (**C^{Ar}**), 129.6 (**C⁹**), 128.5, 127.8 (CH^{Ar}), 125.9 (**C^{Ar}**), 113.7 (**C¹⁰**), 83.0 (**C^{4'}**), 80.9 (**C^{3'}**), 78.2 (**C^{1'}**), 73.6, 71.9 (Bn-CH₂), 71.0 (**C^{5'}**), 55.3 (OCH₃), 53.7 (**C⁷**), 40.7 (**C^{2'}**), 18.9 (CH₃).

NOE (400 MHz, CDCl₃): irradiation at δ 8.09 (H-C(6)) produced NOE enhancement at δ 5.02 (H-C(1')); irradiation at δ 7.59 (H-C(4)) produced NOE enhancements at δ 5.02 (H-C(1')), 2.36 (H-C(CH₃)), 2.05 (H-C(2'_α)), the cross-peak between the H⁴ and the higher field H^{2'_α} resonance signals is stronger than that between H⁴ and H^{1'}; irradiation at δ 5.02 (H-C(1')) produced NOE enhancements at δ 8.09 (H-C(6)), 7.59 (H-C(4)), 3.61 (H-C(5')), 2.61 (H-C(2'_β)), the cross-peak between the H^{1'} and lower field H^{2'_β} resonance signals is very strong, and the cross-peak between H^{1'} and H⁶ resonance signals is stronger than that between H^{1'} and H⁴; irradiation at δ 3.61 (H-C(5')) produced NOE enhancements at δ 5.02 (H-C(1')), 4.60 (H-C(Bn 5')), 4.36 (H-C(4')).

Data for β anomer:

R_f (Hex: EA 1:1, A'): 0.61

LRMS: [ES⁺, MeOH] *m/z* (%): 647 ((M(2 ¹³C)+H)⁺, 9), 646 ((M(¹³C)+H)⁺, 41), 645 ((M+H)⁺, 100).

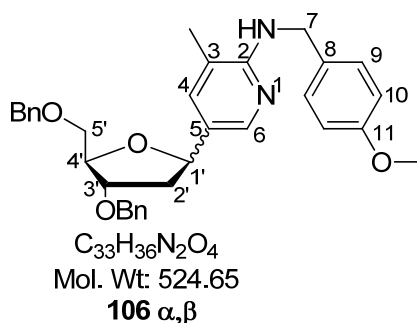
¹H NMR (400 MHz, CDCl₃): δ 8.06 (1H, d, *J* = 2.0 Hz, **H⁶**), 7.40 (1H, s, **H⁴**), 7.29-7.21 (10H, m, **H^{Ar}**), 7.10 (4H, d, *J* = 8.5 Hz, **H⁹**), 6.74 (4H, d, *J* = 9.0 Hz, **H¹⁰**), 5.02 (1H, dd, *J* = 10.5, 4.5 Hz, **H^{1'}**), 4.53 (2H, s, Bn-CH₂), 4.50 (2H, s, Bn-CH₂), 4.23 (1H, m, **H^{4'}**), 4.16-4.12 (5H, m, **H^{3'}** plus **H⁷**), 3.68 (6H, s, OCH₃), 3.60 (1H, dd, *J* = 10.0, 4.0 Hz, **H^{5'}**), 3.53 (1H, dd, *J* = 10.0, 5.0 Hz, **H^{5'}**), 2.27 (4H, m, **H^{2'}_α** plus CH₃), 1.89 (1H, ddd, *J* = 13.0, 11.0, 6.0 Hz, **H^{2'}_β**).

¹³C NMR (100 MHz, CDCl₃): δ 161.3 (**C²**), 158.5 (**C¹¹**), 143.6 (**C⁶**), 138.2, 138.1 (**C^{Ar}**), 137.7 (**C⁴**), 131.1 (**C⁵**), 130.6, 129.5 (**C^{Ar}**), 128.5, 128.4 (**C⁹**), 127.7, 127.7 (**CH^{Ar}**), 127.6 (**C³**), 125.7 (**C^{Ar}**), 113.6, 113.6 (**C¹⁰**), 83.9 (**C^{4'}**), 81.7 (**C^{3'}**), 78.3 (**C^{1'}**), 73.5, 71.2 (Bn-CH₂), 71.1 (**C^{5'}**), 55.2 (OCH₃), 53.7 (**C⁷**), 40.8 (**C^{2'}**), 18.8 (CH₃).

NOE (400 MHz, CDCl₃): irradiation at δ 8.06 (H-C(6)) produced NOE enhancement at δ 5.02 (H-C(1')); irradiation at δ 7.40 (H-C(4)) produced NOE enhancements at δ 5.02 (H-C(1')), 2.27 (H-C(CH₃)), 1.89 (H-C(2'_β)), the cross-peak between H⁴ and higher field H^{2'}_β resonance signals is stronger than that between H⁴ and H^{1'}; irradiation at δ 5.02 (H-C(1')) produced NOE enhancements at δ 8.06 (H-C(6)), 7.40 (H-C(4)), 2.27 (H-C(2'_α)), 1.89 (H-C(2'_β)), the cross peak between H^{1'} and H⁶ resonance signals is stronger than that between H^{1'} and H⁴, and the cross peak between H^{1'} and H^{2'}_α resonance signals is stronger than that between H^{1'} and H^{2'}_β; irradiation at δ 3.60 (H-C(5')) produced NOE enhancements at δ 4.53 (H-C(Bn 5')), 3.53 (H-C(5')), the cross-peak between H^{5'} and H^{5'} resonance signals is stronger than that between H^{5'} and Bn-CH₂; irradiation at δ 3.53 (H-C(5')) produced NOE enhancements at δ 3.60 (H-C(5')), 1.89 (H-C(2'_β)), the cross-peak between H^{5'} and H^{5'} resonance signals is stronger than that between H^{5'} and higher field H^{2'}_β.

Analytical data consistent with that reported in the literature.^{158,202,220}

2-[N-(4-methoxybenzyl)]amino-3-methyl-5-(3',5'-di-O-benzyl-2'-deoxy-D-ribofuranosyl)pyridine, (106 α,β)²²⁰



A solution of **103 RS** (0.79 g, 1.19 mmol) in anhydrous THF (5 mL) was cooled to 0 °C, Bu₃P (0.5 mL, 1.96 mmol) was added followed by DIAD (0.4 mL, 1.96 mmol). The reaction mixture was left to stir at r.t for 17 hrs, after which solvent was removed *in vacuo*. The resulting crude oil in methanol was loaded onto a dowex column. First the reduced DIAD by-product was removed in methanol (500 mL), followed by water (500 mL). The anomeric product was then released from the resin using 10 % NH₃ in MeOH: H₂O 1:1 (4.0 L). The crude, **106 α,β** was taken on to the next step.

Data for α anomer:

LRMS: [ES⁺, MeOH] *m/z* (%): 525 ((M+H)⁺, 100).

Data for β anomer:

LRMS: [ES⁺, MeOH] *m/z* (%): 526 ((M(¹³C)+H)⁺, 38), 525 ((M+H)⁺, 100).

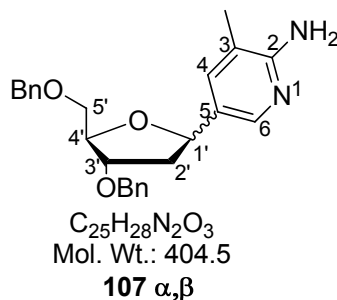
¹H NMR (400 MHz, CDCl₃): δ 8.01 (1H, d, *J* = 2.5 Hz, **H⁶**), 7.36-7.29 (13H, m, **H^{Ar}**, **H⁹** plus **H⁴**), 6.88 (2H, d, *J* = 8.5 Hz, **H¹⁰**), 5.04 (1H, dd, *J* = 11.0, 5.0 Hz, **H^{1'}**), 4.61-4.57 (6H, m, Bn-CH₂ plus **H⁷**), 4.36 (1H, br s, **NH**), 4.27 (1H, td, *J* = 5.0, 2.0 Hz, **H^{4'}**), 4.19 (1H, d, *J* = 6.0 Hz, **H^{3'}**), 3.81 (3H, s, OCH₃), 3.66 (1H, dd, *J* = 10.0, 4.5 Hz, **H^{5'}**), 3.59 (1H, dd, *J* = 10.0, 5.0 Hz, **H^{5'}**), 2.28 (1H, dd, *J* = 12.6, 5.0 Hz, **H^{2'a}**), 2.02 (3H, s, CH₃), 1.96 (1H, ddd, *J* = 13.6, 11.0, 6.0 Hz, **H^{2'β}**).

¹³C NMR (100 MHz, CDCl₃): δ 159.3 (**C²**), 156.8 (**C¹¹**), 143.8 (**C⁶**), 138.7, 138.6 (**C^{Ar}**), 136.1 (**C⁴**), 132.2 (**C^{Ar}**), 129.6 (**C⁹**), 128.9, 128.8, 128.3, 128.2, 128.1, 128.0, 125.7

(CH^{Ar}), 117.2 (C^{Ar}), 114.0 (C¹⁰), 84.1 (C^{4'}), 82.1 (C^{3'}), 79.0 (C^{1'}), 73.9 (Bn-CH₂), 71.7 (C^{5'}), 71.5 (Bn-CH₂), 55.7 (OCH₃), 45.9 (C⁷), 40.9 (C^{2'}), 17.4 (CH₃).

Analytical data consistent with that reported in the literature.^{158,220}

2-Amino-3-methyl-5-(3',5'-di-*O*-benzyl-2'-deoxy-D-ribofuranosyl)pyridine, (107 α,β)²⁰²



Acidic Cleavage of PMB groups Method 1

To a solution of **106 α,β** (0.608 g, 1.16 mmol) in anhydrous DCM (4.0 mL) was added trifluoroacetic acid (4.0 mL) and the solution left to stir for 23 hrs after which solvent was removed *in vacuo*. The residue was redissolved in DCM and quenched with sat. NaHCO₃ (aq). The organic layers were combined and dried over anhydrous sodium sulphate. Column chromatography (DCM: MeOH 1:0, 99:1, 98:2, 97:3 v/v) afforded the mixture **107 α,β** with no separation of the anomers.

Acidic Cleavage of PMB groups Method 2

The crude **104 α,β** (0.055 g, 0.085 mmol) was dissolved in anhydrous DCM (0.5 mL) and trifluoroacetic acid (0.5 mL) was added. The reaction mixture was left to stir at r.t for 19 hrs. On completion, the reaction mixture was quenched with NaHCO₃ (aq) and extracted with DCM. The organic layers were combined and dried over anhydrous sodium sulphate. Solvent was removed *in vacuo* to afford the crude, **107 α,β** .

Data for α anomer:

R_f (DCM: MeOH 8:2, A[']): 0.83

Data for β anomer:

R_f (DCM: MeOH 8:2, A[']): 0.91

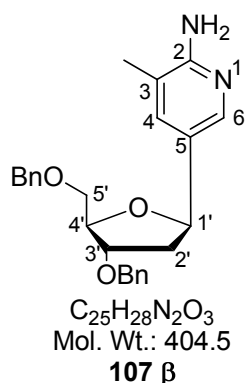
Data for anomeric mixture:

LRMS: [ES⁺, MeOH] *m/z* (%): 809 ((2M+H)⁺, 7), 406 ((M(¹³C)+H)⁺, 28), 405 ((M+H)⁺, 100).

¹H NMR (400 MHz, CDCl₃): δ 7.91 (1H, s, **H⁶ α/β**), 7.89 (1H, d, **H⁶ α/β**), 7.44 (1H, s, **H⁴ α/β**), 7.35-7.28 (21H, m, **H^{Ar} α,β** plus **H⁴ α/β**), 5.02 (1H, dd, $J = 11.0, 5.0$ Hz, **H^{1'} β**), 4.96 (1H, br t, $J = 7.0$ Hz, **H^{1'} α**), 4.62-4.52 (8H, m, **BnCH₂ α,β**), 4.40 (4H, br s, **NH₂ α,β**), 4.34 (1H, dd, $J = 8.5, 4.0$ Hz, **H^{4'} α**), 4.28-4.23 (2H, m, **H^{3'} α** plus **H^{4'} β**), 4.17 (1H, br d, $J = 5.5$ Hz, **H^{3'} β**), 3.65 (1H, dd, $J = 10.5, 4.5$ Hz, **H^{5'} β**), 3.61-3.57 (3H, m, **H^{5'} β** plus **H^{5'} α**), 2.57 (1H, dt, $J = 12.6, 7.0$ Hz, **H^{2'} β α**), 2.28 (1H, dd, $J = 13.6, 5.5$ Hz, **H^{2'} α β**), 2.10 (3H, s, **CH₃ α/β**), 2.06 (3H, s, **CH₃ α/β**), 2.03-2.00 (1H, m, **H^{2'} α α**), 1.94 (1H, ddd, $J = 13.6, 11.0, 6.0$ Hz, **H^{2'} β β**).

¹³C NMR (100 MHz, CDCl₃): δ 157.1, 156.9 (**C² α,β**), 144.1, 144.0 (**C⁶ α,β**), 138.4, 138.2 (**C^{Ar} α,β**), 136.4, 136.3 (**C⁴ α,β**), 128.5, 128.5, 128.5, 127.8, 127.8, 127.7 (**CH^{Ar} α,β**), 83.9 (**C^{4'} β**), 82.9 (**C^{4'} α**), 81.8 (**C^{3'} β**), 80.9 (**C^{3'} α**), 78.5, 78.3 (**C^{1'} α,β**), 73.6, 73.5, 71.9 (**Bn-CH₂ α,β**), 71.3, 71.2 (**C^{5'} α,β**), 71.0 (**Bn-CH₂ α/β**), 40.8, 40.7 (**C^{2'} α,β**), 17.3, 17.2 (**CH₃ α,β**).

2-Amino-3-methyl-5-(3',5'-di-*O*-benzyl-2'-deoxy- β -D-ribofuranosyl)pyridine, (107 β**)²⁰²**



Method 1

To a solution of **104 β** containing reduced DIAD (2.04 g) in anhydrous DCM (14.0 mL), was added trifluoroacetic acid (14.0 mL), and the reaction mixture stirred at r.t for 3 hrs 30 mins. On completion, solvent was removed *in vacuo*. The residue was redissolved in DCM and quenched with sat. $NaHCO_3$ (aq). At this stage the organic phase changed from maroon-brown to green. The organic phases were combined and dried over anhydrous sodium sulphate. Solvent was removed *in vacuo* to give a mustard coloured oil (1.9699 g). Column chromatography (Hex: EA 1:1, 0:1 v/v) afforded **28 β** (0.6289 g, 1.55 mmol) as a mustard yellow oil.

Method 2

To a solution of **104 β** (4.71 g, 7.30 mmol) in anhydrous DCM (17.0 mL) was added trifluoroacetic acid (17.0 mL) and the reaction mixture changed colour from yellow to deep pink-purple, it was left to stir at r.t for 5 hrs 30 mins, before more trifluoroacetic acid (8.0 mL) was added and the reaction mixture left to stir for a further 1 hr 20 mins after which solvent was removed *in vacuo*. The residue was redissolved in water and $NaHCO_3$ was added until > pH 8.0, extraction with DCM followed. The organic layers were combined and dried over anhydrous sodium sulphate. Solvent was removed *in vacuo* to afford the crude (5.1750 g) as a mustard yellow oil. Column chromatography

(Hex: EA 6:4, 1:1, 4:6, 3:7, EA: MeOH 9:1, 8:2 v/v) afforded **107** β (2.92 g, 98.9 %) as a mustard yellow oil.

Method 3

To a solution of **114** β (0.008 g, 0.016 mmol) in anhydrous ethanol (0.6 mL) was added 20 % Pd(OH)₂-C (0.004 g), followed by cyclohexene (0.0002 mL, 0.0016 mmol). The reaction mixture was refluxed (80 °C) for 6 hrs. The reaction mixture was then filtered over celite, and the filtrate was reduced *in vacuo* to give **107** β as a crude residue. (MS confirmed removal of TFA while the benzyl groups remained intact).

R_f (EA: MeOH 95:5, A'): 0.43

LRMS: [ES⁺, MeOH] *m/z* (%): 831 ((2M+Na)⁺, 10), 427 ((M+Na)⁺, 12), 406 ((M(C¹³)+H)⁺, 28), 405 ((M+H)⁺, 100).

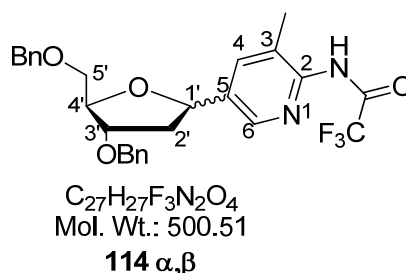
HRMS: [ES⁺] Found 405.2175 Da (M+H)⁺, calculated 405.2173 Da.

¹H NMR (400 MHz, CDCl₃): δ 7.81 (1H, s, **H**⁶), 7.41 (1H, s, **H**⁴), 7.35-7.29 (10H, m, **H**^{Ar}), 5.50 (2H, br s, **NH**₂), 5.00 (1H, dd, *J* = 10.6, 5.0 Hz, **H**^{1'}), 4.58 (2H, s, Bn-CH₂), 4.56 (2H, s, Bn-CH₂), 4.28-4.25 (1H, m, **H**^{4'}), 4.18 (1H, d, *J* = 6.0 Hz, **H**^{3'}), 3.64 (1H, dd, *J* = 10.0, 4.0 Hz, **H**^{5'}), 3.58 (1H, dd, *J* = 10.0, 5.0 Hz, **H**^{5'}), 2.29 (1H, dd, *J* = 13.0, 5.0 Hz, **H**^{2''}_a), 2.08 (3H, s, **CH**₃), 1.91 (1H, ddd, *J* = 13.0, 10.6, 5.5 Hz, **H**^{2''} _{β}).

¹³C NMR (100 MHz, CDCl₃): δ 156.2 (**C**²), 139.8 (**C**⁶) 138.2, 138.1 (**C**^{Ar}), 137.8 (**C**⁴), 128.6, 128.5, 127.9, 127.8, 127.8, 127.7 (**CH**^{Ar}), 127.1 (**C**⁵), 118.4 (**C**³), 84.0 (**C**^{4'}), 81.7 (**C**^{3'}), 78.0 (**C**^{1'}), 73.6 (Bn-CH₂), 71.2 (Bn-CH₂), 71.2 (**C**^{5'}), 40.7 (**C**^{2'}), 16.9 (**CH**₃).

Analytical data consistent with that reported in the literature.^{158,202,220}

3-Methyl-2-trifluoroacetamide-5-(3',5'-di-*O*-benzyl-2'-deoxy-D-ribofuranosyl)pyridine, (114** α,β)**²⁵⁴



107 α,β (0.2967 g, 0.733 mmol) was dissolved in anhydrous pyridine (3 mL) and cooled to 0 °C. Trifluoroacetic anhydride (0.21 mL) was then added and the reaction mixture left to stir at r.t for 72 hrs. Solvent was removed *in vacuo*. The residue was quenched with sat. KCl and extracted with DCM. The organic phases were combined and dried over anhydrous sodium sulphate. Solvent was removed *in vacuo*. Column chromatography (Hex: EA 1:0, 9:1, 8:2, 7:3 v/v) afforded the starting material **107** α,β (0.2117 g) and an insignificant quantity of product **114** α,β .

Data for both anomers:

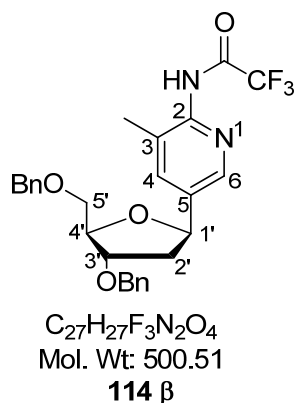
Data for α anomer:

R_f (Hex: EA 1:1, A'): 0.37

Data for β anomer:

R_f (Hex: EA 1:1, A'): 0.48

3-Methyl-2-trifluoroacetamide-5-(3',5'-di-*O*-benzyl-2'-deoxy- β -D-ribofuranosyl)-pyridine, (114** β)**^{251,254}



A solution of **107** β (1.58 g, 3.91 mmol) in anhydrous pyridine (17.0 mL) was cooled to 0 °C before dropwise addition of trifluoroacetic anhydride (0.81 mL, 5.86 mmol). The reaction mixture was then left to stir at r.t for 2 hrs 30 mins, after which the solvent volume was reduced to half under high vacuum and the reaction mixture was partitioned between sat. $NaHCO_3$ (aq) and DCM. The organic phases were combined and dried over anhydrous sodium sulphate; solvent was removed *in vacuo* to afford **114** β (1.36 g, 70 %) as a yellow oil.

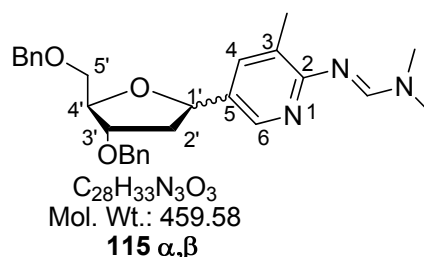
LRMS: $[ES^+, MeOH]$ m/z (%): 1024 ((2M(^{13}C)+Na), 4), 1023 ((2M(^{13}C)+Na), 28), 523 ((M + Na) $^+$, 100), 501 ((M + H) $^+$, 22).

1H NMR (400 MHz, $CDCl_3$): δ 8.02 (1H, s, H^6), 7.71 (1H, s, H^4), 7.37-7.30 (10H, m, CH^{Ar}), 5.12 (1H, dd, $J = 10.5, 5.0$ Hz, $H^{1'}$), 4.58-4.56 (4H, m, Bn- CH_2), 4.31 (1H, td, $J = 4.0, 2.0$ Hz, $H^{4'}$), 4.20 (1H, br d, $J = 6.0$ Hz, $H^{3'}$), 3.64 (1H, dd, $J = 10.0, 6.0$ Hz, $H^{5'}$), 3.62 (1H, dd, $J = 10.0, 5.5$ Hz, $H^{5'}$), 2.38 (1H, ddd, $J = 13.0, 5.0, 1.0$ Hz, $H^{2'a}$), 2.26 (3H, s, CH_3), 1.91 (1H, ddd, $J = 13.0, 10.5, 5.5$ Hz, $H^{2'a}$).

^{13}C NMR (100 MHz, $CDCl_3$): δ 149.0 (C^2), 139.4 (C^4), 138.1 (C^6), 138.0 (C^{Ar}), 136.0 (C^5), 131.0 (C^3), 128.7, 128.6, 128.0, 127.8, 127.7 (CH^{Ar}), 84.4 ($C^{4'}$), 81.6 ($C^{3'}$), 77.7 ($C^{1'}$), 73.7 (Bn- CH_2), 71.4 (Bn- CH_2), 71.1 ($C^{5'}$), 41.3 ($C^{2'}$), 17.7 (CH_3).

^{19}F NMR (300 MHz, $CDCl_3$): δ -75.7 (s, CF_3).

2-*N,N*-[(dimethylamino)methylidene]-3-methyl-5-(3',5'-di-*O*-benzyl-2'-deoxy-D-ribofuranosyl)pyridine, (115** α,β)**²⁵⁵



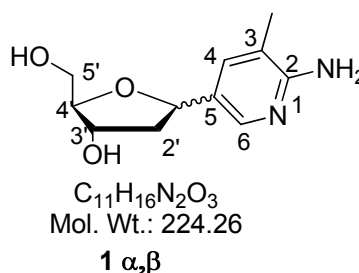
To a solution of **107** α,β (0.2117 g, 0.523 mmol) in anhydrous methanol (2.0 mL), was added *N,N*-dimethylformamide dimethylacetal (0.7 mL, 5.23 mmol) and the reaction mixture stirred at 70 °C for 26 hrs. After which solvent was removed *in vacuo*. The crude, **115** α,β (0.2492 g) was taken onto the next step.

Data for anomeric mixture:

R_f (DCM: MeOH 95:5, A'): 0.52

LRMS: [ES⁺, MeOH] *m/z* (%): 461 ((M(¹³C)+H)⁺, 30), 460 ((M+H)⁺, 100).

2-Amino-3-methyl-5-(2'-deoxy-D-ribofuranosyl)pyridine, (1** α,β)**



De-benzylation Method 1²⁷¹

To a solution of **115** α,β (0.2492 g, 0.54 mmol) in anhydrous methanol (1.5 mL) was added 20 % Pd(OH)₂-C (0.4227 g), followed by formic acid (1.5 mL). The reaction mixture was heated to 50 °C for 6 hrs. The reaction mixture was then filtered over

celite, the filtrate was reduced *in vacuo*. The residue was quenched with water and NaOH and solvent removed *in vacuo*. The residue was stirred in ammonia solution at r.t overnight. Solvent was then removed *in vacuo*. The crude was loaded onto dowex using methanol to remove the salt formed. The dowex was washed with methanol then water, to remove the salt. The anomeric product was released from the dowex using 10 % NH₃ in MeOH: H₂O 1:1, solvent was reduced *in vacuo* and under high vacuum to give crude (0.1231 g). Column chromatography (EA: MeOH 7:3 v/v) was attempted, however, no separation of the anomers was achieved. NMR confirmed amidine protection had been removed in conjunction with the benzyl groups, affording **1** α,β .

De-benzylation Method 2¹⁵⁸

A solution of **107** β (0.305 g, 0.75 mmol) in anhydrous DCM (4.0 mL) was cooled to - 78 °C, before careful addition of boron tribromide solution (1.0 M in DCM, 2.64 mL, 2.64 mmol). The dark brown reaction mixture was left to stir at - 78 °C for 1 hr 30 mins after which anhydrous methanol (4.5 mL) was added and the reaction mixture was left to stir at - 60 °C for 1 hr 30 mins. The reaction mixture was then left to warm to r.t for 1 hr 30 mins. Solvent was then removed *in vacuo* and the residue redissolved in DCM and extracted with water. The aqueous layers were combined and neutralised with ammonia solution until pH 7.0. Some solvent was removed *in vacuo* and the aqueous phase was stored at 4 °C to allow for precipitation. Precipitation did not occur; hence, solvent was completely removed *in vacuo* and the crude purified on dowex to remove any salts. NMR was messy and showed that the product contained anomers, **1** α,β ($\alpha:\beta$ 1:2).

De-benzylation Method 3²⁷⁷

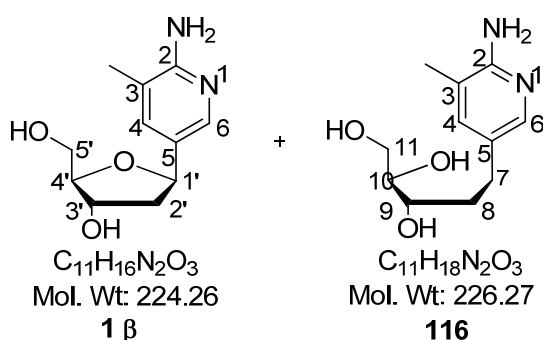
A solution of **107** α (0.3478 g, 0.86 mmol) in anhydrous DCM (4.0 mL) was cooled to - 78 °C before dropwise addition of boron trichloride (1.0 M in DCM, 4.3 mL, 4.3 mmol). The reaction mixture was left to stir at - 78 °C for 3 hrs 30 mins after which methanol (3.7 mL) was added and the reaction mixture left to warm to r.t over 3 hrs after which NaHCO₃ was added until pH 7.0 and water added. Partition between DCM and water followed. The aqueous phases were combined and reduced *in vacuo* to give the crude as a white solid plus salt. NMR and MS analyses showed that the benzyl

groups had been removed, however, the ring had opened and re-closed to form an anomeric mixture of the nucleoside **1** α,β ($\alpha:\beta$ 1:2).

R_f (DCM: MeOH 8:2, A'): 0.35

LRMS: [ES⁺, MeOH] *m/z* (%): 226 ((M(C¹³)+H)⁺, 12), 225 ((M+H)⁺, 100).

2-Amino-3-methyl-5-(2'-deoxy- β -D-ribofuranosyl)pyridine, (1** β); 2-Amino-3-methyl-5-(1',2'-deoxy-D-ribofuranosyl)pyridine, (**116**)**



To a solution of **107** β (0.6283 g, 1.553 mmol) in anhydrous ethanol (1.5 mL) was added Pd(OH)₂-C (0.20 g). The reaction mixture was put under high vacuum followed by H₂ purge (x 3); it was left to stir under H₂ gas and heated to 60 °C for 80 hrs. The reaction mixture was allowed to cool to r.t and then filtered over celite. Solvent from the filtrate was removed *in vacuo* to give a faint orange oil (0.4056 g). Column chromatography (EA:MeOH 98:2, 97:3, 96:4, 95:5 v/v) afforded **1** β (0.088g, 25.1 %) as a clear oil, and an open chain by-product, **116**, as was confirmed by NMR and MS.

Data for 1 β :

R_f (EA: MeOH 8:2, A'): 0.16

LRMS: [ES⁺, MeOH] *m/z* (%): 471 ((2M+Na)⁺, 14), 247 ((M+Na)⁺, 28), 226 ((M(C¹³)+H)⁺, 12), 225 ((M+H)⁺, 100).

HRMS: [ES⁺] Found 225.1232 Da (M+H)⁺, calculated 225.1234 Da.

¹H NMR (400 MHz, DMSO): δ 7.72 (1H, d, J = 2.0 Hz, **H⁶**), 7.24 (1H, s, **H⁴**), 5.68 (2H, br s, **NH₂**), 4.98 (1H, d, J = 4.0 Hz, **OH^{3'}**), 4.81 (1H, dd, J = 10.5, 5.5 Hz, **H^{1'}**), 4.69 (1H, br t, J = 5.5 Hz, **OH^{5'}**), 4.15 (1H, m, **H^{3'}**), 3.69 (1H, td, J = 5.5, 2.0 Hz, **H^{4'}**), 3.41 (2H, m, **H^{5'}**), 2.03 (3H, s, **CH₃**), 1.91 (1H, dd, J = 12.6, 5.0 Hz, **H^{2'}_a**), 1.80 (1H, td, J = 10.6, 5.5 Hz, **H^{2'}_b**).

¹³C NMR (100 MHz, DMSO): δ 156.8 (**C²**), 141.0 (**C⁶**), 136.8 (**C⁴**), 125.8 (**C⁵**), 117.0 (**C³**), 87.6 (**C^{4'}**), 77.0 (**C^{1'}**), 72.4 (**C^{3'}**), 62.4 (**C^{5'}**), 42.7 (**C^{2'}**), 16.9 (**CH₃**).

Analytical data consistent with that stated in the literature.²⁰²

Data for 116:

R_f (EA: MeOH 8:2, A[']): 0.04

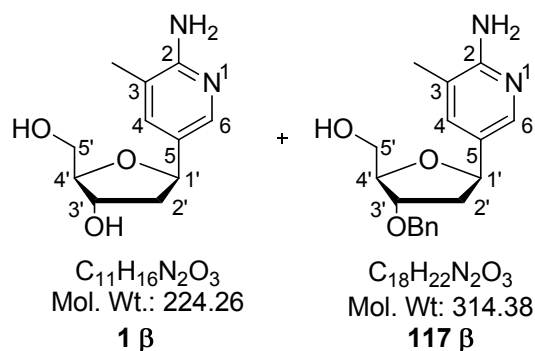
LRMS: [ES⁺, MeOH] m/z (%): 228 ((M(¹³C)+H)⁺, 12), 227 ((M+H)⁺, 100).

HRMS: [ES⁺] Found 227.1392 Da (M+H)⁺, calculated 227.1390 Da.

¹H NMR (400 MHz, DMSO): δ 7.61 (1H, s, **H⁶**), 7.11 (1H, s, **H⁴**), 5.51 (2H, br s, **NH₂**), 4.48 (3H, br s, **OH**), 3.52 (1H, dd, J = 11.0, 3.5 Hz, **H¹¹**), 3.33 (1H, dd, J = 10.5, 6.0 Hz, **H¹¹**), 3.30-3.22 (2H, m, **H⁹** plus **H¹⁰**), 2.56 (1H, ddd, J = 14.0, 10.0, 4.5 Hz, **H⁷**), 2.36 (1H, ddd, J = 16.6, 9.5, 7.0 Hz, **H⁷**), 2.02 (3H, s, **CH₃**), 1.80-1.73 (1H, m, **H⁸**), 1.47 (1H, tt, J = 13.5, 5.0 Hz, **H⁸**).

¹³C NMR (100 MHz, DMSO): δ 156.1 (**C²**), 143.4 (**C⁶**), 137.8 (**C⁴**), 126.1 (**C⁵**), 115.9 (**C³**), 74.8 (**C⁹**), 70.5 (**C¹⁰**), 63.5 (**C¹¹**), 34.9 (**C⁸**), 27.5 (**C⁷**), 17.0 (**CH₃**).

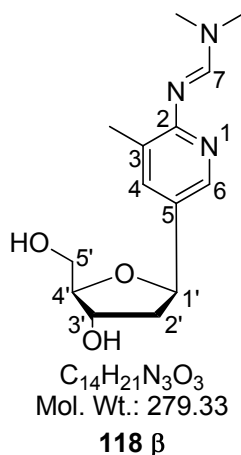
2-Amino-3-methyl-5-(2'-deoxy-β-D-ribofuranosyl)pyridine, (1 β); 2-Amino-3-methyl-5-(3'-O-benzyl-2'-deoxy-β-D-ribofuranosyl)pyridine, (117 β)



To a solution of **107 β** (0.827 g, 2.04 mmol) in anhydrous ethanol (13.0 mL) was added $Pd(OH)_2 \cdot C$ (0.25 g). The reaction mixture was put under high vacuum followed by H_2 purge (x 3); it was left under H_2 gas and heated to 60 °C for 21 hrs 30 mins. After which all starting material had disappeared, product and intermediate were present. The reaction mixture was allowed to cool to r.t and then filtered over celite. Solvent from the filtrate was removed *in vacuo* to give a faint pink foam (0.5884 g). Column chromatography (EA:MeOH 98:2, 97:3, 96:4, 95:5 v/v) afforded **1 β** (0.095 g, 0.423 mmol) as a white foam. The intermediate, **117 β** (0.22 g, 0.7 mmol) was also isolated and put on for further hydrogenolysis reaction (31 hrs) to yield **1 β** (0.045 g, 0.201 mmol). Overall **1 β** (0.14 g, 31 %) was afforded as a white foam.

Analytical data for **1 β** consistent with that reported above and stated in the literature.²⁰²

2-*N,N*-[(dimethylamino)methylidene]-3-methyl-5-(2'-deoxy- β -D-ribofuranosyl)pyridine, (118 β**)²⁵⁵**



Method 1²⁵⁵

To a solution of **1 α,β** (0.184 g, 0.808 mmol) in anhydrous methanol (4.0 mL), was added *N,N*-dimethylformamide dimethyl acetal (3.0 mL, excess) and the solution stirred at reflux (70 °C) for 4 hrs 30 mins, after which reaction mixture was cooled to r.t and solvent removed *in vacuo*. The crude (0.2684 g) still contained a little of the starting material. Initially, column chromatography (DCM: MeOH 1:0, 99:1, 97:3, 95:5 v/v) separated the starting material from the anomeric product. Another column was attempted to separate the anomers (EA: H-1 95:5, 9:1 v/v), however separation was not achieved. A third column (EA: H-1 1:0, 99:1, 98:2 v/v) afforded **118 β** (0.02 g) as a light yellow oil and anomeric mixture **118 α,β** (0.0327 g) as a mustard yellow oil. Overall yield 23.4 %.

Method 2²⁵⁵

To a solution of **1 β** (0.095 g, 0.42 mmol) in anhydrous methanol (1.5 mL) was added *N,N*-dimethylformamide dimethyl acetal (1.5 mL, excess) and the reaction mixture stirred at reflux (70 °C) for 20 hrs. On completion, solvent was removed *in vacuo* to give a yellow residue. Column chromatography (DCM: MeOH 1:0, 95:5, 9:1, 85:15 v/v) afforded **118 β** (> 0.071 g, > 61 %) as a yellow oil.

R_f (DCM: MeOH 8:2, A'): 0.47

LRMS: [ES⁺, MeOH] *m/z* (%): 281 ((M(C¹³)+H)⁺, 17), 280 ((M+H)⁺, 100).

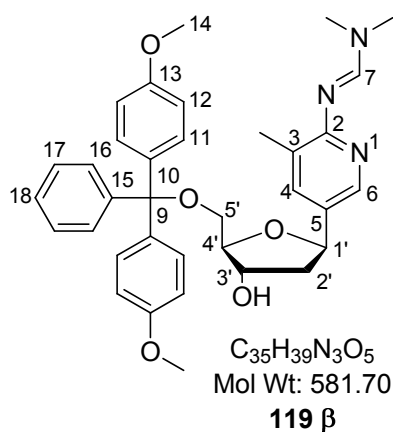
HRMS: [ES⁺] Found 280.1650 Da (M+H)⁺, calculated 280.1656 Da.

¹H NMR (400 MHz, MeOD): δ 7.90 (1H, s, **H**⁷), 7.85 (1H, d, *J* = 2.0 Hz, **H**⁶), 7.42 (1H, d, *J* = 1.5 Hz, **H**⁴), 4.91 (1H, dd, *J* = 10.5, 5.5 Hz, **H**^{1'}), 4.19 (1H, m, **H**^{3'}), 3.77 (1H, td, *J* = 5.0, 2.5 Hz, **H**^{4'}), 3.51 (2H, m, **H**^{5'}), 2.97 (3H, s, NCH₃), 2.93 (3H, s, NCH₃), 2.12 (3H, s, CH₃), 2.01 (1H, dd, *J* = 13.0, 5.5 Hz, **H**^{2'}_α), 1.83 (1H, ddd, *J* = 13.0, 10.6, 6.0 Hz, **H**^{2'}_β).

¹³C NMR (100 MHz, MeOD): δ 161.9 (**C**²), 156.5 (**C**⁷), 144.0 (**C**⁶), 138.0 (**C**⁴), 132.8 (**C**⁵), 127.9 (**C**³), 89.2 (**C**^{4'}), 79.3 (**C**^{1'}), 74.4 (**C**^{3'}), 64.0 (**C**^{5'}), 44.5 (**C**^{2'}), 40.8, 34.8 (NCH₃), 17.9 (CH₃).

Analytical data consistent with that reported in the literature.²⁵⁵

2-*N,N*-[(dimethylamino)methylidene]-3-methyl-5-(5'-*O*-(4,4'-dimethoxytrityl)-2'-deoxy-β-D-ribofuranosyl)pyridine, (119 β**)**¹⁵⁸



To a solution of **118 β** (0.3336g, 1.20 mmol; which had been co-evaporated with anhydrous pyridine (2 mL x 3) under high vacuum and left to dry overnight under high vacuum) in anhydrous pyridine (2 mL) was added a solution of 4,4'-dimethoxytrityl chloride (0.51 g, 1.51 mmol) in anhydrous pyridine (3 mL) dropwise over 20 mins. The

reaction mixture was left to stir at r.t for 2 hrs 20 mins after which solvent was reduced to half under high vacuum. TEA (0.5 mL) and methanol (1 mL) were added, the colour of the reaction mixture changed from bright orange to green. The reaction mixture was partitioned between NaHCO₃ and DCM. The aqueous layer was further extracted with DCM (x 2). The organic phases were combined and dried over anhydrous sodium sulphate. Solvent was removed *in vacuo* to give crude (0.7874 g) as a yellow foam. Column chromatography (DCM: MeOH 1:0, 95.5:0.5, 99:1, 98.5:1.5, 98:2, 97.5:2.5, 97:3, 96.5:3.5, 96:4 v/v (all pre-equilibrated with 0.5 % pyridine)) afforded **119** β (0.435 g, 62.4 %) as a white foam.

R_f (DCM: MeOH 9:1, A'): 0.46

LRMS: [ES⁺, MeOH] *m/z* (%): 583 ((M(C¹³)+H)⁺, 28), 582 ((M+H)⁺, 100).

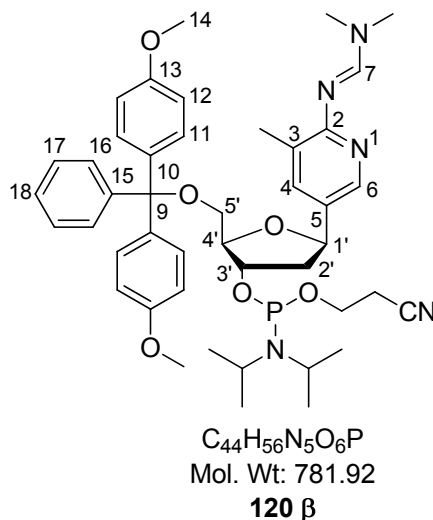
HRMS: [ES⁺] Found 582.2966 Da (M+H)⁺, calculated 582.2962 Da.

¹H NMR (400 MHz, CDCl₃): δ 8.29 (1H, s, **H**⁷), 7.99 (1H, d, *J* = 2.0 Hz, **H**⁶), 7.43 (1H, d, *J* = 1.5 Hz, **H**⁴), 7.41 (2H, br s, **H**¹⁶), 7.31 (4H, dd, *J* = 9.0, 2.0 Hz, **H**¹¹), 7.23 (2H, br s, **H**¹⁷), 7.17 (1H, br t, *J* = 7.5 Hz, **H**¹⁸), 6.78 (4H, d, *J* = 9.0 Hz, **H**¹²), 5.05 (1H, dd, *J* = 10.0, 5.5 Hz, **H**^{1'}), 4.41 (1H, m, **H**^{3'}), 3.99 (1H, td, *J* = 5.5, 2.5 Hz, **H**^{4'}), 3.79 (6H, s, OCH₃), 3.31 (1H, dd, *J* = 10.0, 4.5 Hz, **H**^{5'}), 3.22 (1H, dd, *J* = 10.0, 5.5 Hz, **H**^{5'}), 3.04 (6H, s, NCH₃), 2.20 (3H, s, CH₃), 2.15 (1H, ddd, *J* = 13.6, 6.0, 2.0 Hz, **H**^{2'a}), 2.04 (1H, ddd, *J* = 13.0, 10.6, 6.0 Hz, **H**^{2'p}).

¹³C NMR (100 MHz, CDCl₃): δ 160.5 (**C**²), 158.6 (**C**¹³), 154.5 (**C**⁷), 145.0 (**C**¹⁵), 143.6 (**C**⁶), 136.2 (**C**¹⁰), 136.2 (**C**⁴), 130.5 (**C**⁵), 130.2 (**C**¹¹), 128.3, 128.0 (**C**¹⁶/**C**¹⁷), 126.9 (**C**¹⁸), 126.5 (**C**³), 113.3 (**C**¹²), 86.4 (**C**^{4'}), 86.4 (**C**⁹), 78.2 (**C**^{1'}), 75.1 (**C**^{3'}), 64.7 (**C**^{5'}), 55.4 (OCH₃), 43.7 (**C**^{2'}), 40.6, 34.6 (NCH₃) 17.9 (CH₃).

Analytical data consistent with that reported in the literature.¹⁵⁸

2-*N,N*-[(dimethylamino)methylidene]-3-methyl-5-(5'-*O*-(4,4'-dimethoxytrityl)-2'-deoxy- β -D-ribofuranosyl)pyridine-3'-*O*-(2-cyanoethyl-*N,N*-diisopropyl)phosphoramidite, (120 β**)²⁷⁵**



To a solution of **119 β** (0.380 g, 0.65 mmol) in anhydrous DCM (3 mL) was added *N,N*-diisopropylethylamine (0.23 mL, 1.31 mmol) followed by dropwise addition of 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite (0.19 mL, 0.85 mmol) under an argon atmosphere and the reaction mixture left to stir at r.t for 1 hr 45 mins. The reaction mixture was transferred *via* needle and syringe to a separating funnel containing anhydrous DCM. The reaction mixture was washed with saturated KCl (10 mL) and the organic phase was dried over anhydrous sodium sulphate. Solvent was removed *in vacuo* under argon, to give following purification by column chromatography (EA: DCM 8:2 (0.5 % pyridine)) under argon pressure, **120 β** , as a white foam (diastereoisomers 1:1, 0.4014 g, 79.1 %).

R_f (EA (0.5 % pyridine), A'): 0.40, 0.31.

LRMS: [ES⁺, MeOH] *m/z* (%): 784 ((M(2 ¹³C)+H)⁺, 10), 783 ((M(¹³C)+H)⁺, 43), 782 ((M+H)⁺, 100).

HRMS: [ES⁺] Found 782.4044 Da (M+H)⁺, calculated 782.4041 Da.

¹H NMR (400 MHz, CDCl₃): δ 8.31 (2H, s, **H⁷**), 8.05 (2H, br t, *J* = 3.0 Hz, **H⁶**), 7.48-7.45 (6H, m, **H⁴** plus **H¹⁶/H¹⁷**), 7.34 (8H, dt, *J* = 9.0, 3.0 Hz, **H¹¹**), 7.28-7.23 (4H, m,

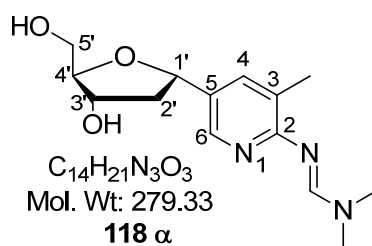
$\text{H}^{16}/\text{H}^{17}$), 7.21-7.17 (2H, m, H^{18}), 6.81-6.78 (8H, m, H^{12}), 5.07 (2H, dd, $J = 10.5, 5.0$ Hz, $\text{H}^{1'}$), 4.55-4.51 (2H, m, $\text{H}^{3'}$), 4.23-4.15 (2H, m, $\text{H}^{4'}$), 3.87-3.79 (2H, m, $\text{OCH}_2\text{CH}_2\text{CN}$), 3.77 (12H, s, OCH_3), 3.73-3.65 (2H, m, $\text{OCH}_2\text{CH}_2\text{CN}$), 3.64-3.53 (4H, m, iPr-CH), 3.32-3.20 (4H, m, $\text{H}^{5'}$), 3.05 (12H, s, NCH_3), 2.61 (2H, t, $J = 8.0$ Hz, $\text{OCH}_2\text{CH}_2\text{CN}$), 2.46 (2H, t, $J = 8.0$ Hz, $\text{OCH}_2\text{CH}_2\text{CN}$), 2.31 (2H, ddd, $J = 32.0, 12.0, 4.0$ Hz, $\text{H}^{2'_{\alpha}}$), 2.22 (6H, s, CH_3), 2.11-2.03 (2H, m, $\text{H}^{2'_{\beta}}$), 1.29-1.08 (24H, m, iPr- CH_3).

^{13}C NMR (100 MHz, CDCl_3): δ 158.6 (C^2), 154.4 (C^7), 145.1, 145.0 (C^6), 136.3 (C^4), 130.5 (C^5), 130.3, 130.3 (C^{11}), 128.5, 128.4, 127.9 ($\text{C}^{16}/\text{C}^{17}$), 126.9 (C^3), 126.9 (C^{18}), 113.2 (C^{12}), 86.3 ($\text{C}^{4'}$), 86.1 (C^9), 78.6 ($\text{C}^{1'}$), 76.3 ($\text{C}^{3'}$), 64.4 ($\text{C}^{5'}$), 58.6, 58.4 ($\text{OCH}_2\text{CH}_2\text{CN}$), 55.4 (OCH_3), 45.5, 43.5, 43.4, 43.3, (iPr-CH), 43.2 ($\text{C}^{2'}$), 38.0, 35.5 (NCH_3), 24.7, 24.6, 24.6, 23.1 (iPr- CH_3), 20.5 ($\text{OCH}_2\text{CH}_2\text{CN}$), 17.8 (CH_3).

^{31}P NMR (121 MHz, CDCl_3) δ 148.69 (s, 1P, **P**), 148.49 (s, 1P, **P**).

Analytical data consistent with that reported in the literature.¹⁵⁸

2-*N,N*-[(dimethylamino)methylidene]-3-methyl-5-(2'-deoxy- α -D-ribofuranosyl)pyridine, (118 α**)**²⁵⁵



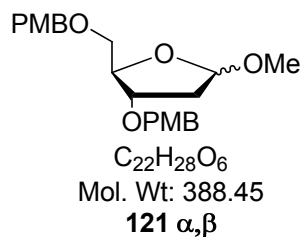
To a solution of **1 α** (0.0787 g, 0.35 mmol) in anhydrous methanol (1.6 mL) was added *N,N*-dimethylformamide dimethylacetal (1.6 mL) and the reaction mixture stirred at 70 °C for 25 hrs 30 mins. On completion, solvent was removed *in vacuo* to give crude **118 α** (0.1984 g) as a yellow residue.

R_f (DCM: MeOH 8:2, A⁺): 0.35

LRMS: [ES^+ , MeOH] m/z (%): 281 (($\text{M}^{(13}\text{C})+\text{H}$)⁺, 19), 280 (($\text{M}+\text{H}$)⁺, 100).

¹H NMR (300 MHz, MeOD): δ 8.01 (1H, s, H⁷), 7.96 (1H, s, H⁶), 7.54 (1H, s, H⁴), 4.97 (1H, dd, *J* = 8.7, 6.4 Hz, H^{1'}), 4.32 (1H, m, H^{3'}), 3.95 (1H, m, H^{4'}), 3.65 (2H, m, H^{5'}), 3.08 (3H, s, NCH₃), 3.03 (3H, s, NCH₃), 2.59 (1H, dt, *J* = 12.4, 6.8 Hz, H^{2'β}), 2.24 (3H, s, CH₃), 1.88 (1H, ddd, *J* = 12.5, 9.1, 7.2 Hz, H^{2'α}).

Methyl-3',5'-di-*O*-paramethoxybenzyl-2'-deoxy-D-ribose, (121)^{202,253}



To a solution of 2'-deoxy-D-ribose (1.0 g, 7.46 mmol) in anhydrous methanol (37 mL), a solution of acetyl chloride (0.03 mL, 0.45 mmol) in anhydrous methanol (2.24 mL) was added dropwise and the reaction mixture left to stir at r.t for 3 hrs 20 mins after which sat. NaHCO₃ aq. was added until pH 7.0. The reaction mixture was filtered and solvent removed *in vacuo*. The residue was co-evaporated with toluene (x 2). The residue was redissolved in anhydrous DMF (37 mL), *p*-methoxybenzyl chloride (4 mL, 29.8 mmol) was added and the reaction mixture cooled to 0 °C before portion wise addition of NaH (60 %, 1.2 g, 29.8 mmol). The reaction mixture was left to warm to r.t and stirred for 15 hrs. On completion, the reaction mixture was cooled to 0 °C before careful addition of methanol (27 mL). Solvent was removed *in vacuo* and co-evaporated with toluene. The residue was then quenched with water and extracted with DCM. The organic phases were combined and dried over anhydrous sodium sulphate. Solvent was removed *in vacuo*. Purification by column chromatography (DCM: EA 1:0, 98:2, 9:1, 8:2, 7:3, 6:4 v/v) afforded **121 α,β** (2.7387 g, 98 %) as a faint yellow oil.

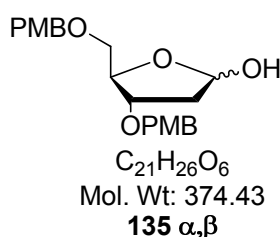
R_f (DCM: EA 99:1, A'): 0.10.

¹H NMR (400 MHz, CDCl₃): δ 7.37-7.25 (8H, m, H^{Ar} α,β), 6.94-6.88 (8H, m, H^{Ar} α,β), 5.13 (1H, dd, *J* = 5.0, 2.0, H^{1'} α/β), 5.10 (1H, dd, *J* = 5.5, 1.0, H^{1'} α/β), 4.56-4.44 (8H, m, PMB-CH₂ α,β), 4.30-4.24 (2H, m, H^{4'} α,β), 4.18-4.14 (1H, m, H^{3'} α/β), 4.01-3.97 (1H, m, H^{3'} α/β), 3.85 (6H, s, PMB-OCH₃ α/β), 3.84 (6H, s, PMB-OCH₃ α/β), 3.59-

3.50 (4H, m, $\text{H}^{5'}$ α,β), 3.45 (3H, s, OCH_3 α/β), 3.35 (3H, s, OCH_3 α/β), 2.29-2.22 (2H, m, $\text{H}^{2'}$ α,β), 2.15 (1H, dt, $J = 13.5, 5.5$, $\text{H}^{2'}$ α/β), 2.03 (1H, m, $\text{H}^{2'}$ α/β).

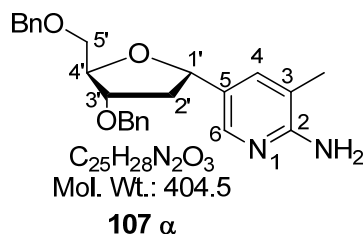
^{13}C NMR (100 MHz, CDCl_3): δ 159.3, 159.2 (C^{Ar} α,β), 129.5, 129.4, 129.3, 113.8, 113.8, 113.7 (CH^{Ar} α,β), 105.5, 105.2 ($\text{C}^{1'}$), 82.9, 82.2 ($\text{C}^{4'}$), 86.1 (C^9), 79.7, 78.2 ($\text{C}^{3'}$), 73.1, 73.0 (PMB-CH_2 α,β), 71.8 ($\text{C}^{5'}$ α/β), 71.3, 71.2 (PMB-CH_2 α,β), 69.9 ($\text{C}^{5'}$ α/β), 55.4 (PMB-OCH_3 α,β), 55.1, 55.0 (OCH_3 α,β), 39.4, 38.9 ($\text{C}^{2'}$ α,β).

3',5'-Di-*O*-paramethoxybenzyl-2'-deoxy-D-ribose, (135)^{261,262}



A solution of **121 α,β** (2.6559, 7.09 mmol) in 80 % acetic acid solution (24 mL acetic acid, 6 mL water) was heated to 110 °C for 1 hr 30 mins after which reaction mixture was allowed to cool to r.t. sat. NaHCO_3 was added until pH 7.0, extraction with DCM followed. The organic phases were combined and dried over anhydrous sodium sulphate. Solvent was removed *in vacuo* to give crude (1.9712 g) as a brown oil.

2-Amino-3-methyl-5-(3',5'-di-*O*-benzyl-2'-deoxy- α -D-ribofuranosyl)pyridine, (107 α)²⁰²



To a solution of **104 α** (2.71 g, 4.20 mmol) in anhydrous DCM (10 mL) was added trifluoroacetic acid (10 mL) over 5 mins (the reaction mixture changed colour from yellow to deep pink-purple), it was left to stir at r.t for 4 hrs 45 mins after which solvent

was removed *in vacuo*. The residue was redissolved in water and sodium carbonate was added until pH 9.0, extraction with DCM followed. The organic layers were combined and dried over anhydrous sodium sulphate. Solvent was removed *in vacuo* to afford the crude (2.66 g) as a mustard yellow oil. Purification by column chromatography (Hex: EA 6:4, 1:1, 0:1, EA: MeOH 8:2 v/v) afforded **107 a** (1.51 g, 89 %) as a yellow oil.

R_f (DCM: MeOH 95:5, B'): 0.19.

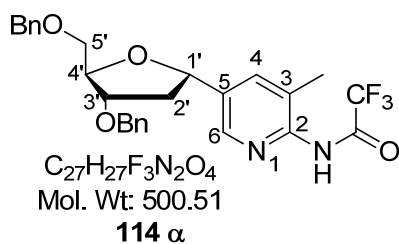
LRMS: [ES⁺, MeOH] *m/z* (%): 832 ((2M(¹³C)+Na)⁺, 7), 831 ((2M+Na)⁺, 12), 810 ((2M(C¹³)+H)⁺, 9), 809 ((2M+H)⁺, 19), 427 ((M+Na)⁺, 12), 406 ((M(C¹³)+H)⁺, 30), 405 ((M+H)⁺, 100).

HRMS: [ES⁺] Found 405.2175 Da (M+H)⁺, calculated 405.2173 Da.

¹H NMR (400 MHz, CDCl₃): δ 7.89 (1H, s, **H**⁶), 7.44 (1H, d, *J* = 1.0 Hz, **H**⁴), 7.35-7.28 (10H, m, **H**^{Ar}), 4.96 (1H, dd, *J* = 8.6, 7.0 Hz, **H**^{1'}), 4.62-4.47 (4H, m, Bn-CH₂), 4.45 (2H, br s, NH₂), 4.34 (1H, dd, *J* = 8.5, 4.5 Hz, **H**^{4'}), 4.25 (1H, dd, *J* = 10.5, 6.5 Hz, **H**^{3'}), 3.60 (2H, m, **H**^{5'}), 2.57 (1H, dt, *J* = 12.6, 7.0 Hz, **H**^{2'β}), 2.10 (3H, s, CH₃), 2.07-2.00 (1H, m, **H**^{2'α}).

¹³C NMR (100 MHz, CDCl₃): δ 156.9 (**C**²), 143.9 (**C**⁶), 138.2 (**C**⁵), 136.4 (**C**⁴), 128.5, 127.8, 127.7 (CH^{Ar}), 116.8 (**C**³), 82.8 (**C**^{4'}), 80.9 (**C**^{3'}), 78.3 (**C**^{1'}), 73.6, 71.8 (Bn-CH₂), 70.9 (**C**^{5'}), 40.7 (**C**^{2'}), 17.2 (CH₃).

3-Methyl-2-trifluoroacetamide-5-(3',5'-di-*O*-benzyl-2'-deoxy- α -D-ribofuranosyl)pyridine, (114 a**)²⁵¹**



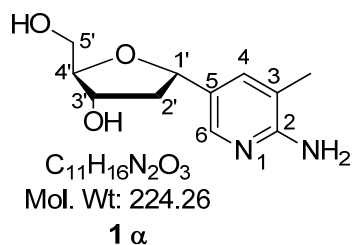
A solution of **107 a** (1.04 g, 2.57 mmol) in anhydrous pyridine (11.0 mL) was cooled to 0 °C before dropwise addition of trifluoroacetic anhydride (0.53 mL, 3.85 mmol). The reaction mixture was then left to stir at r.t for 2 hrs 45 mins, after which the reaction mixture was partitioned between sat. $NaHCO_3$ (aq) and DCM. The organic layers were combined and dried over anhydrous sodium sulphate; solvent was reduced *in vacuo* to afford **114 a** (1.25 g, 97 %) as a mustard-yellow oil.

R_f (DCM: MeOH 95:5, A'): 0.77.

LRMS: $[ES^+, MeOH]$ m/z (%): 1024 ($(2M(^{13}C) + Na)^+$, 9), 1023 ($(2M(^{13}C) + Na)^+$, 90), 524 ($(M(^{13}C + Na)^+$, 5), 523 ($(M + Na)^+$, 100), 502 ($(M(^{13}C) + H)^+$, 20), 501 ($(M + H)^+$, 75).

^{19}F NMR (300 MHz, $CDCl_3$): δ -76.0 (s, CF_3).

2-Amino-3-methyl-5-(2'-deoxy- α -D-ribofuranosyl)pyridine, (1 a**)²⁷⁷**



A solution of **114 a** (1.25 g, 2.5 mmol) in anhydrous DCM (7.0 mL) was cooled to - 78 °C before dropwise addition of boron trichloride (1.0 M in DCM, 7.5 mL, 7.5 mmol). The reaction mixture was left to stir at - 78 °C for 7 hrs after which methanol (30 mL) was added and the reaction mixture left to stand at 4 °C for 18 hrs.

NaOH (aq, 2M) was added until pH 7.0. Solvent was removed *in vacuo*. The crude was partitioned between water and DCM. The aqueous phases were combined and solvent reduced *in vacuo* and high vacuum, to give crude (1.88 g) as a cream-yellow solid. Column chromatography (DCM: MeOH 85:15, 8:2 v/v) afforded **1 a** (0.54 g, 96 %) as a cream-white solid.

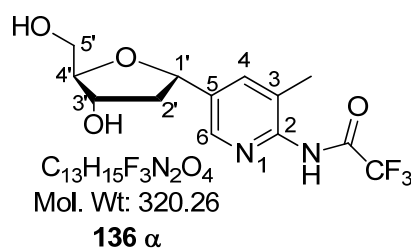
R_f (DCM: MeOH 8:2, B'): 0.14

LRMS: [ES⁺, MeOH] *m/z* (%): 225 ((M + H)⁺, 100).

¹H NMR (400 MHz, DMSO): δ 7.74 (1H, d, *J* = 2.0 Hz, **H⁶**), 7.57 (1H, s, **H⁴**), 6.74 (2H, br s, **NH₂**), 5.06 (1H, br s, **OH^{3'}**), 4.82 (1H, t, *J* = 7.5 Hz, **H^{1'}**), 4.70 (1H, br s, **OH^{5'}**), 4.18 (1H, m, **H^{3'}**), 3.80 (1H, dd, *J* = 9.5, 5.0 Hz, **H^{4'}**), 3.48 (1H, dd, *J* = 11.6, 4.0 Hz, **H^{5'}**), 3.40 (1H, dd, *J* = 11.6, 5.0 Hz, **H^{5'}**), 2.44 (1H, dt, *J* = 12.6, 7.0 Hz, **H^{2'p}**), 2.11 (3H, s, **CH₃**), 1.74 (1H, ddd, *J* = 12.6, 8.5, 6.5 Hz, **H^{2'a}**).

¹³C NMR (100 MHz, DMSO): δ 155.1 (**C²**), 138.4 (**C⁴**), 136.8 (**C⁶**), 127.2 (**C⁵**), 118.6 (**C³**), 86.8 (**C^{4'}**), 76.5 (**C^{1'}**), 72.0 (**C^{3'}**), 62.1 (**C^{5'}**), 42.8 (**C^{2'}**), 17.1 (**CH₃**).

3-Methyl-2-trifluoroacetamide-5-(2'-deoxy-α-D-ribofuranosyl)pyridine, (136 α)²⁵¹



A solution of **1 a** (0.307 g, 1.371 mmol) in anhydrous pyridine (5.7 mL) was cooled to 0 °C before dropwise addition of trifluoroacetic anhydride (0.3 mL, 2.16 mmol). The reaction mixture was left to stir at r.t for 3 hrs after which the reaction mixture was partitioned between DCM and NaHCO₃. The organic phases were combined and dried over anhydrous sodium sulphate. Solvent was then removed *in vacuo* and high vacuum. The residue was dried under high vacuum. The product, **136 α** was present in a mixture with **1 a**.

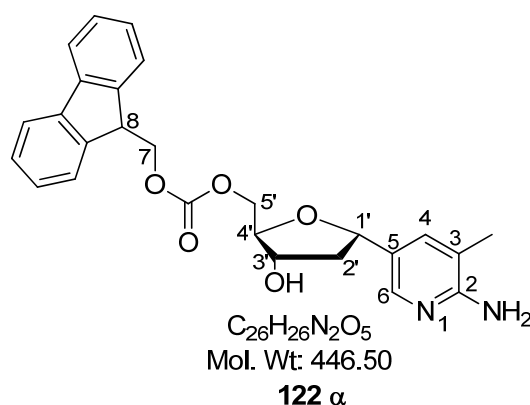
R_f (DCM: MeOH 8:2, A'): 0.54

LRMS: [ES⁺, MeOH] *m/z* (%): 664 ((2M(C¹³) + Na)⁺, 21), 663 ((M(C¹³) + Na)⁺, 95), 343 ((M + Na)⁺, 100).

¹H NMR (400 MHz, MeOD): δ 8.78 (1H, br s, NH), 7.89 (1H, s, H⁶), 7.72 (1H, s, H⁴), 5.03 (1H, t, *J* = 7.0 Hz, H^{1'}), 4.40-4.35 (1H, m, H^{3'}), 4.01 (1H, dd, *J* = 8.5, 4.0 Hz, H^{4'}), 3.75-3.58 (2H, m, H^{5'}), 2.65 (1H, dt, *J* = 13.0, 6.5 Hz, H^{2'β}), 2.27 (3H, s, CH₃), 1.89 (1H, dt, *J* = 13.0, 6.5 Hz, H^{2'α}).

¹⁹F NMR (300 MHz, MeOD): δ -78.2, -78.1, -77.9, -77.8 (s, CF₃).

2-Amino-3-methyl-5-(5'-*O*-(fluorenylmethoxycarbonyl)-2'-deoxy-α-D-ribofuranosyl)pyridine, (122 α)



To a solution of **1 α** (0.1 g, 0.45 mmol) in anhydrous pyridine (1.0 mL) was added a solution of Fmoc-Cl (0.583 g, 2.25 mmol) in anhydrous pyridine (1.0 mL). The reaction mixture was left to stir at r.t for 7 hrs after which the reaction mixture was partitioned between EA and water. The organic phases were combined and dried over anhydrous sodium sulphate. Solvent was removed *in vacuo* to give crude (0.2792 g) as a white solid. Column chromatography (DCM: MeOH 1:0, 98:2, 95:5, 9:1 v/v) afforded **122 α**.

R_f (DCM: MeOH 8:2, A⁺): 0.71

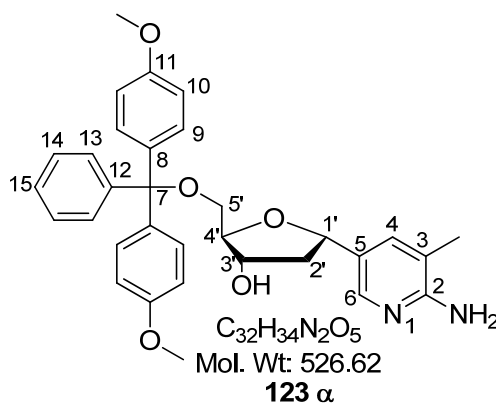
LRMS: [ES⁺, MeOH] *m/z* (%): 893 ((2M+H)⁺, 4), 470 ((M(¹³C) + Na)⁺, 7), 269 ((M + Na)⁺, 24), 448 ((M(¹³C) + H)⁺, 29), 447 ((M+H)⁺, 100).

¹H NMR (400 MHz, DMSO): δ 7.89 (2H, d, *J* = 7.6 Hz, Fmoc H^{Ar}), 7.71 (1H, s, H⁶), 7.66 (2H, d, *J* = 7.5 Hz, Fmoc H^{Ar}), 7.43-7.31 (5H, m, Fmoc H^{Ar} plus H⁴), 6.04 (2H, br

s, NH), 5.28 (1H, br s, OH^{3'}), 4.76 (1H, dd, $J = 9.5, 6.5$ Hz, H^{1'}), 4.54 (2H, d, $J = 6.0$ Hz, H⁷), 4.31 (1H, t, $J = 6.0$ Hz, H⁸), 4.21 (1H, dd, $J = 11.5, 3.0$ Hz, H^{5'}), 4.15-4.07 (2H, m, H^{5'} plus H^{3'}), 3.94 (1H, td, $J = 7.0, 3.5$ Hz, H^{4'}), 2.42 (1H, dt, $J = 12.5, 6.5$ Hz, H^{2'}_β), 2.06 (3H, s, CH₃), 1.77 (1H, ddd, $J = 12.0, 9.0, 7.6$ Hz, H^{2'}_α).

¹³C NMR (100 MHz, DMSO): δ 157.1 (Fmoc C=O), 154.5 (C²), 143.3, 143.3 (Fmoc C^{Ar}), 141.2 (C⁶), 140.8 (Fmoc C^{Ar}), 136.5 (C⁴), 127.7, 127.2 (Fmoc CH^{Ar}), 126.1 (C⁵), 124.9 (Fmoc CH^{Ar}), 121.4 (Fmoc C^{Ar}), 120.2 (Fmoc CH^{Ar}), 120.0 (Fmoc C^{Ar}), 116.8 (C³), 82.2 (C^{4'}), 76.6 (C^{1'}), 71.7 (C^{3'}), 68.7 (C⁷), 67.8 (C^{5'}), 46.3 (C⁸), 42.2 (C^{2'}), 18.5 (CH₃).

2-Amino-3-methyl-5-(5'-O-(4,4'-dimethoxytrityl)-2'-deoxy-α-D-ribofuranosyl)pyridine, (123 α)¹⁵⁸



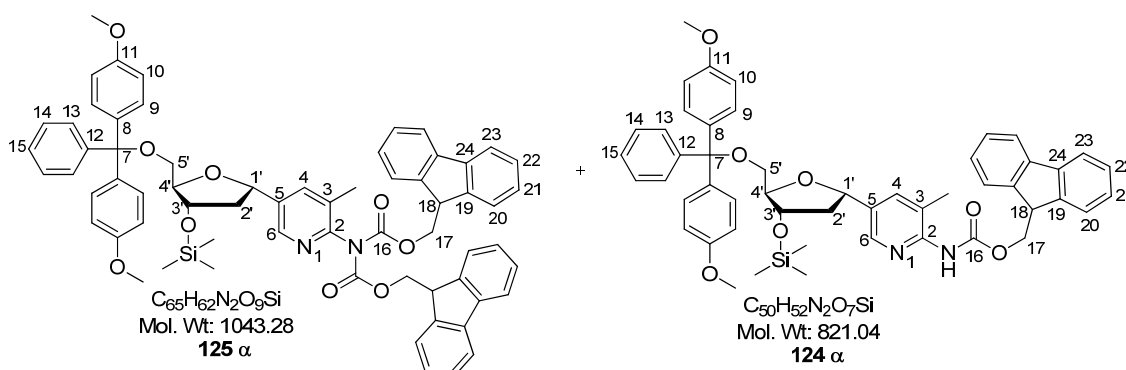
To a solution of **1 α** (0.24 g, 1.10 mmol) in anhydrous pyridine (5.0 mL) was added a solution of 4,4'-dimethoxytrityl chloride (0.47 g, 1.39 mmol) in anhydrous pyridine (5.0 mL) and the reaction mixture left to stir at r.t for 3 hrs before solvent volume was reduced to half under high vacuum. TEA (0.5 mL) and methanol (1mL) were added and the reaction mixture partitioned between DCM and sat. NaHCO₃ (aq). The organic phases were combined and dried over anhydrous sodium sulphate. Solvent was removed *in vacuo* and under high vacuum to give crude as a yellow foam (0.4293 g). Column chromatography (DCM: MeOH (0.5 % pyridine) 1:0, 99:1, 98:2, 97:3, 96:4, 95:5 v/v) afforded **123 α** (0.1347 g, 23 %) as a pale yellow foam.

LRMS: [ES⁺, MeOH] m/z (%): 1054 ((2M(¹³C) + H)⁺, 8), 1053 ((2M + H)⁺, 19), 528 ((M(¹³C) + H)⁺, 28), 527 ((M + H)⁺, 100).

¹H NMR (400 MHz, DMSO): δ 7.80-7.74 (1H, m, **H**⁶), 7.43-7.36 (3H, m, **H**⁴, **H**¹³/**H**¹⁴), 7.30-7.27 (6H, m, **H**⁹ plus **H**¹³/**H**¹⁴), 7.22 (1H, m, **H**¹⁵), 6.88 (4H, d, *J* = 9.0 Hz, **H**¹⁰), 5.66 (2H, br s, **NH**₂), 5.07 (1H, d, *J* = 5.0 Hz, **OH**^{3'}), 4.78 (1H, dd, *J* = 9.6, 6.0 Hz, **H**^{1'}), 4.17-4.14 (1H, m, **H**^{3'}), 3.95 (1H, dd, *J* = 9.0, 5.5 Hz, **H**^{4'}), 3.74 (6H, s, **OCH**₃), 3.06 (1H, dd, *J* = 10.0, 4.0 Hz, **H**^{5'}), 3.02 (1H, dd, *J* = 11.0, 5.5 Hz, **H**^{5'}), 2.41 (1H, dt, *J* = 12.0, 6.6 Hz, **H**^{2'}_β), 2.06 (3H, s, **CH**₃), 1.77 (1H, ddd, *J* = 12.0, 9.5, 7.0 Hz, **H**^{2'}_α).

¹³C NMR (100 MHz, DMSO): δ 158.0 (**C**²), 158.0 (**C**¹¹), 149.6 (**C**^{Ar}), 145.0 (**C**⁶), 136.1 (**C**⁴), 135.8 (**C**⁸), 135.8, 135.5 (**C**^{Ar}), 129.7, 127.8, 127.7 (**C**⁹, **C**¹³, **C**¹⁴), 126.6 (**C**¹⁵), 113.1 (**C**¹⁰), 85.2 (**C**⁷), 84.1 (**C**^{4'}), 77.0 (**C**^{1'}), 72.3 (**C**^{3'}), 64.3 (**C**^{5'}), 55.0 (**OCH**₃), 42.8 (**C**^{2'}), 17.1 (**CH**₃).

2-[*N,N*-bis-(fluorenylmethyloxycarbonyl)]amino-3-methyl-5-(3'-trimethylsilyl-5'-*O*-(4,4'-dimethoxytrityl)-2'-deoxy- α -D-ribofuranosyl)pyridine, (125 α) and 2-[*N*-(fluorenylmethyloxycarbonyl)]amino-3-methyl-5-(3'-trimethylsilyl-5'-*O*-(4,4'-dimethoxytrityl)-2'-deoxy- α -D-ribofuranosyl)pyridine, (124 α)



To a solution of **123 α** (0.06 g, 0.114 mmol) in anhydrous pyridine (1.0 mL) was added TMS-Cl (0.1 mL, 0.798 mmol) and the reaction mixture left to stir at r.t for 1 hr 50 mins after which another 7 eq TMS-Cl (0.1 mL, 0.798 mmol) was added and the reaction mixture left to stir at r.t for 40 mins. Fmoc-Cl (0.15 g, 0.57 mmol) was then added, and the reaction mixture went from a clear yellow solution to opaque. The reaction mixture was left to stir at r.t for 30 mins after which water was added and the reaction mixture left to stir at r.t for 1 hr 30 mins, extraction with DCM followed. The organic phases were combined and dried over anhydrous sodium sulphate. Solvent was removed *in vacuo* and high vacuum to give crude as a yellow oil (0.1544 g). Column

chromatography (Pet ether: EA (0.5 % Pyridine) 9:1, 85:15, 8:2, 7:3, 6:4 v/v) afforded a mixture of **125 α** and **124 α** (0.0458, 48.7 %) as a white solid.

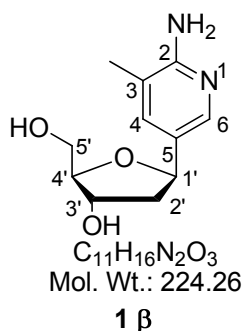
Data for 125 α:

LRMS: [ES⁺, MeOH] *m/z* (%): 1067 ((M(¹³C) + Na⁺, 22), 1066 ((M + Na)⁺, 80).

Data for 124 α:

LRMS: [ES⁺, MeOH] *m/z* (%): 845 ((M(C¹³)+Na)⁺, 12), 844 ((M+Na)⁺, 22), 821 ((M + H)⁺, 100).

2-Amino-3-methyl-5-(2'-deoxy-β-D-ribofuranosyl)pyridine, (1 β)²⁷⁷

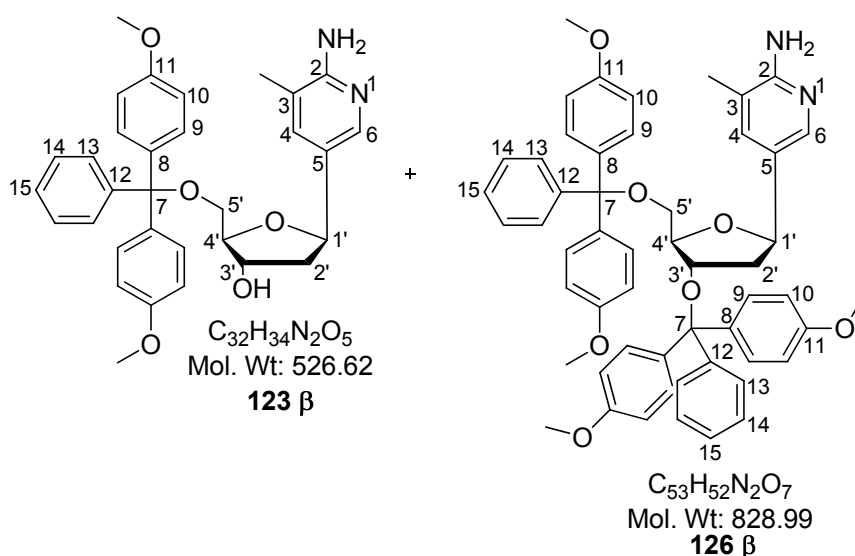


A solution of **114 β** (1.36 g, 2.72 mmol) in anhydrous DCM (8.0 mL) was cooled to – 78 °C before dropwise addition of boron trichloride (1.0 M in DCM, 8.2 mL, 8.2 mmol). The reaction mixture was left to stir at - 78 °C for 4 hrs 30 mins after which excess methanol (30 mL) was added and the reaction mixture left to stand at 4 °C for 15 hrs 30 mins after which NaOH (aq, 2.0 M) was added until pH 8.0. Solvent was reduced *in vacuo*. The residue was partitioned between water and DCM. The aqueous phases were combined and solvent removed *in vacuo*, to give crude as a cream-yellow solid. Column chromatography (DCM: MeOH 85:15, 8:2 v/v) afforded **1 β** (> 0.60 g, > 98.5 %) as a cream-white solid.

R_f (DCM: MeOH 8:2, A⁺): 0.25

Analytical data consistent with that reported above and stated in the literature.²⁰²

2-Amino-3-methyl-5-(5'-O-(4,4'-dimethoxytrityl)-2'-deoxy- β -D-ribofuranosyl)pyridine, (123 β) and 2-amino-3-methyl-5-(3',5'-di-O-(4,4'-dimethoxytrityl)-2'-deoxy- β -D-ribofuranosyl)pyridine, (126 β)¹⁵⁸



To a solution of **1 β** (0.66 g, 2.95 mmol) in anhydrous pyridine (6.0 mL) was added a solution of 4,4'-dimethoxytrityl chloride (1.30 g, 3.83 mmol) in anhydrous pyridine (7.0 mL) drop wise over 15 mins and the reaction mixture left to stir at r.t for 2 hrs after which solvent volume was reduced to half *in vacuo*. TEA (1.0 mL) and MeOH (2.0 mL) were added, the reaction mixture became a lime green colour from bright orange. The reaction mixture was partitioned between DCM and sat. $NaHCO_3$ (aq). The aqueous layers were further extracted with DCM. The organic phases were combined and dried over anhydrous sodium sulphate. Solvent was removed *in vacuo* and high vacuum to give crude (1.63 g) as a yellow foam. Column chromatography (DCM: MeOH (0.5 % Pyridine) 1:0, 99:1, 97:3, 95:5 v/v) afforded **123 β** (0.4596 g) and **126 β** (0.2451 g) as white foams (39.5 % overall yield).

Data for 123 β :

R_f (DCM: MeOH 8:2, A'): 0.43

LRMS: [ES⁺, MeOH] *m/z* (%): 1054 ((2M(¹³C) + H)⁺, 8), 1053 ((2M + H)⁺, 18), 528 ((M(¹³C) + H)⁺, 28), 527 ((M + H)⁺, 100).

HRMS: [ES⁺] Found 527.2539 Da (M+H)⁺, calculated 527.2540 Da.

¹H NMR (400 MHz, DMSO): δ 7.77 (1H, s, **H**⁶), 7.42 (2H, d, *J* = 7.5 Hz, **H**¹³/**H**¹⁴), 7.32-7.19 (7H, m, **H**⁴, **H**⁹, **H**¹³/**H**¹⁴), 7.16 (1H, m, **H**¹⁵), 6.87 (4H, d, *J* = 9.0 Hz, **H**¹⁰), 5.64 (2H, br s, **NH**₂), 5.06 (1H, d, *J* = 4.0 Hz, **OH**^{3'}), 4.88 (1H, dd, *J* = 10.0, 5.0 Hz, **H**^{1'}), 4.15 (1H, m, **H**^{3'}), 3.86 (1H, dd, *J* = 7.0, 4.5 Hz, **H**^{4'}), 3.73 (6H, s, **OCH**₃), 3.06 (2H, m, **H**^{5'}), 1.99 (1H, ddd, *J* = 13.0, 5.5, 1.5 Hz, **H**^{2'}_a), 1.96 (3H, s, **CH**₃), 1.88 (1H, ddd, *J* = 12.0, 10.0, 6.0 Hz, **H**^{2'}_β).

¹³C NMR (100 MHz, DMSO): δ 158.0 (**C**²), 158.0 (**C**¹¹), 145.0 (**C**⁶), 143.8 (**C**^{Ar}), 135.7 (**CH**^{Ar}), 135.5 (**C**⁴), 129.7, 128.9 (**C**^{Ar}), 128.2, 127.8, 127.7, 126.6 (**C**⁹, **C**¹³, **C**¹⁴, **C**¹⁵), 125.4 (**C**⁵), 115.5 (**C**³), 113.1 (**C**¹⁰), 85.9 (**C**^{4'}), 85.3 (**C**⁷), 77.5 (**C**^{1'}), 72.5 (**C**^{3'}), 64.9 (**C**^{5'}), 55.0 (**OCH**₃), 43.2 (**C**^{2'}), 17.0 (**CH**₃).

Data for 126 β:

R_f (DCM: MeOH 8:2, A'): 0.89

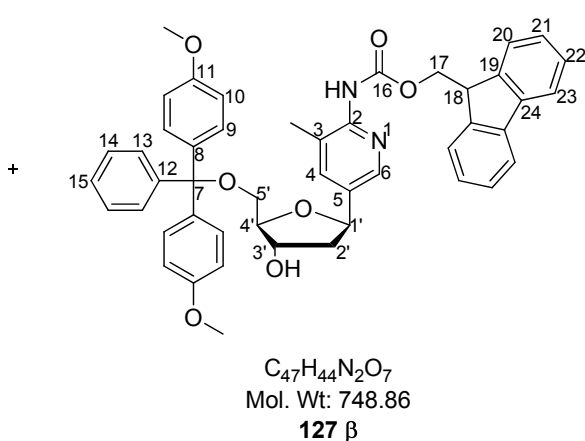
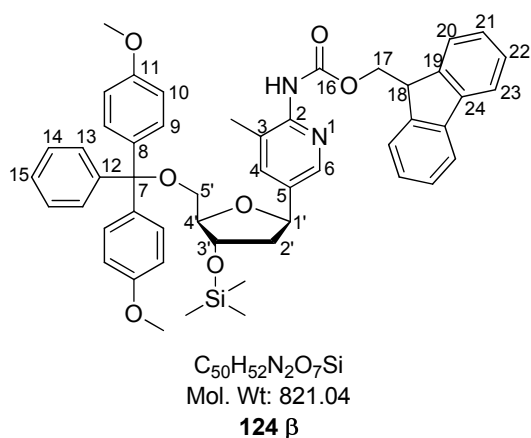
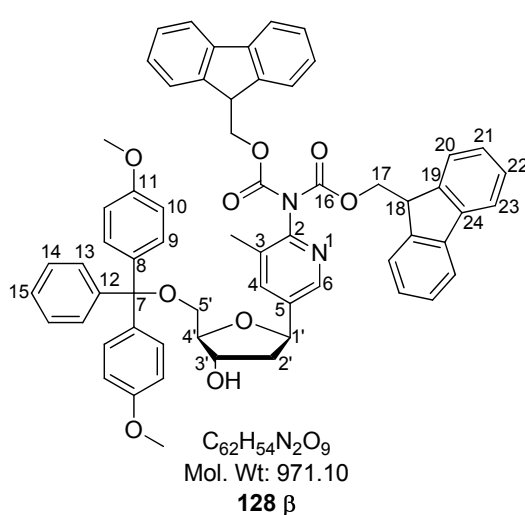
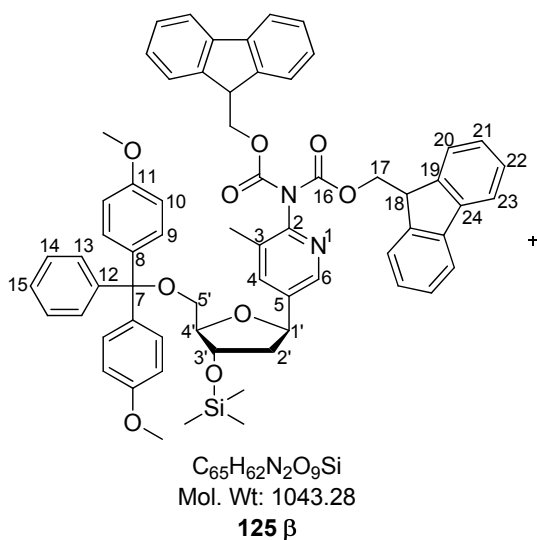
LRMS: [ES⁺, MeOH] *m/z* (%): 1182 ((2M + H)⁺, 4), 830 ((M(¹³C) + H)⁺, 46), 829 ((M + H)⁺, 100).

HRMS: [ES⁺] Found 829.3847 Da (M+H)⁺, calculated 829.3847 Da.

¹H NMR (400 MHz, DMSO): δ 7.73 (1H, s, **H**⁶), 7.37-7.35 (2H, m, **H**^{Ar}), 7.29-7.13 (17H, m, **H**^{Ar}), 6.83-6.75 (8H, m, **H**¹⁰), 5.67 (2H, s, **NH**₂), 4.88 (1H, dd, *J* = 11.0, 5.0 Hz, **H**^{1'}), 4.16 (1H, d, *J* = 5.0 Hz, **H**^{3'}), 3.91 (1H, s, **H**^{4'}), 3.73 (6H, s, **OCH**₃), 3.71 (6H, s, **OCH**₃), 2.94-2.91 (1H, m, **H**^{5'}), 2.78 (1H, dd, *J* = 10.0, 4.5 Hz, **H**^{5'}), 1.90 (3H, s, **CH**₃), 1.59-1.51 (1H, m, **H**^{2'}_a), 1.37 (1H, dd, *J* = 12.6, 5.0 Hz, **H**^{2'}_β).

¹³C NMR (100 MHz, DMSO): δ 158.0 (**C**²), 158.0 (**C**¹¹), 157.9, 157.9 (**C**¹¹), 145.3, 144.8 (**C**^{Ar}), 143.3 (**C**⁶), 136.1 (**C**^{Ar}), 135.5, 135.5, 135.2, 129.8, 129.7, 129.6, 129.5, 127.8, 127.8, 127.6, 127.6, 126.7, 126.5 (**CH**^{Ar}), 113.1, 113.0 (**C**¹⁰), 86.2 (**C**^{4'}), 85.2, 84.4 (**C**⁷), 77.8 (**C**^{1'}), 76.3 (**C**^{3'}), 64.2 (**C**^{5'}), 54.9 (**OCH**₃), 42.0 (**C**^{2'}), 16.9 (**CH**₃).

2-[*N,N*-bis-(fluorenylmethyloxycarbonyl)]amino-3-methyl-5-(3'-trimethylsilyl-5'-*O*-(4,4'-dimethoxytrityl)-2'-deoxy- β -D-ribofuranosyl)pyridine, (125 β); 2-[*N,N*-bis-(fluorenylmethyloxycarbonyl)]amino-3-methyl-5-(5'-*O*-(4,4'-dimethoxytrityl)-2'-deoxy- β -D-ribofuranosyl)pyridine, (128 β); 2-[*N*-(fluorenylmethyloxycarbonyl)]amino-3-methyl-5-(3'-trimethylsilyl-5'-*O*-(4,4'-dimethoxytrityl)-2'-deoxy- β -D-ribofuranosyl)pyridine, (124 β) and 2-[*N*-(fluorenylmethyloxycarbonyl)]amino-3-methyl-5-(5'-*O*-(4,4'-dimethoxytrityl)-2'-deoxy- β -D-ribofuranosyl)pyridine, (127 β)



To a solution of **123 β** (0.3972 g, 0.755 mmol) in anhydrous pyridine (4.0 mL) was added TMS-Cl (0.57 g, 4.53 mmol) and the reaction mixture left to stir at r.t for 2 hrs after which another 3 eq TMS-Cl (0.3 mL, 2.36 mmol) was added and the reaction mixture left to stir at r.t for 1 hr. Another 3 eq TMS-Cl (0.3 mL, 2.36 mmol) was added and the reaction mixture left to stir at r.t for 45 mins after which another 3 eq TMS-Cl (0.3 mL, 2.36 mmol) added and the reaction mixture left to stir at r.t for 45 mins, solvent was then removed in high vacuum. The residue was resuspended in anhydrous pyridine (4.0 mL) and Fmoc-Cl (0.78 g, 3.02 mmol) added portion wise after which white solid precipitated. The reaction mixture was left to stir at r.t for 1 hr 20 mins after which water (excess) was added and the reaction mixture left to stir at r.t for 1 hr. Partition between DCM and water followed. The organic phases were combined and dried over anhydrous sodium sulphate. Solvent was removed *in vacuo* and high vacuum to give crude (0.9532 g) as a peach foam. Column chromatography (Pet ether: EA (0.5 % Pyridine) 1:0, 95:5, 9:1, 85:15, 8:2, 7:3, 6:4, 1:1, EA:MeOH (0.5 % Pyridine) 9:1 v/v) afforded **125 β** (0.2056 g, 0.2 mmol), **128 β** (0.147 g, 0.18 mmol), **124 β** (0.125 g, 0.13 mmol) and **127 β** (0.083 g, 0.11 mmol); all as white foams (overall yield 82.1 %).

Data for 125 β:

R_f (Pet ether: EA 6:4, A'): 0.73

LRMS: [ES⁺, MeOH] *m/z* (%): 1069 ((M(3 ¹³C) + Na)⁺, 12), 1068 ((M(2 ¹³C) + Na)⁺, 35), 1067 ((M(¹³C) + Na)⁺, 80), 1066 ((M + Na)⁺, 100).

¹H NMR (400 MHz, DMSO): δ 8.17 (1H, d, *J* = 2.0 Hz, **H⁶**), 7.74 (4H, d, *J* = 7.0 Hz, **H^{Ar}**), 7.59 (1H, d, *J* = 2.0 Hz, **H⁴**), 7.46-7.40 (2H, m, **H^{Ar}**), 7.32-7.13 (19H, m, **H^{Ar}**), 6.84 (4H, d, *J* = 9.0 Hz, **H¹⁰**), 5.19 (1H, dd, *J* = 10.0, 6.0 Hz, **H¹**), 4.50-4.37 (5H, m, **H¹⁷** plus **H³**), 4.07 (2H, t, *J* = 6.0 Hz, **H¹⁸**), 4.01 (1H, dd, *J* = 7.5, 5.0 Hz, **H⁴**), 3.67 (6H, s, OCH₃), 3.27-3.20 (2H, m, **H⁵**), 2.27-2.24 (1H, m, **H^{2'}**_w), 1.96 (1H, ddd, *J* = 14.0, 10.0, 6.0 Hz, **H²**_β), 1.68 (3H, s, CH₃), 0.08 (9H, s, TMS-CH₃).

¹³C NMR (100 MHz, DMSO): δ 158.1, 150.6, 148.2, 144.8 (**C^{Ar}**), 144.2 (**C⁶**), 143.7, 143.1, 140.6, 138.3 (**C^{Ar}**), 137.4 (**C⁴**), 135.6, 130.5 (**C^{Ar}**), 129.7, 127.6, 126.9, 124.7,

120.0 (CH^{Ar}), 113.2 (C^{10}), 86.2 ($\text{C}^{4'}$), 85.7 (C^7), 76.8 ($\text{C}^{1'}$), 73.7 ($\text{C}^{3'}$), 68.0 (C^{17}), 63.8 ($\text{C}^{5'}$), 55.0 (OCH_3), 45.9 (C^{18}), 43.4 ($\text{C}^{2'}$), 16.0 (CH_3), 0.0 (TMS-CH_3).

Data for 128 β :

R_f (Pet ether: EA 6:4, A'): 0.27

LRMS: $[\text{ES}^+, \text{MeOH}] m/z$ (%): 995 ($(\text{M}(2\ ^{13}\text{C}) + \text{Na})^+$, 3), 994 ($(\text{M}(^{13}\text{C}) + \text{Na})^+$, 60), 993 ($(\text{M} + \text{Na})^+$, 100).

HRMS: $[\text{ES}^+]$ Found 993.3737 Da ($\text{M} + \text{Na})^+$, calculated 993.3722 Da.

^1H NMR (400 MHz, DMSO): δ 8.15 (1H, d, $J = 2.0$ Hz, H^6), 7.74 (4H, d, $J = 7.5$ Hz, H^{Ar}), 7.58 (1H, d, $J = 2.0$ Hz, H^4), 7.46-7.44 (2H, m, H^{Ar}), 7.35-7.13 (19H, m, H^{Ar}), 6.83 (4H, d, $J = 8.5$ Hz, H^{10}), 5.27 (1H, d, $J = 4.0$ Hz, $\text{OH}^{3'}$), 5.20 (1H, dd, $J = 10.0, 5.5$ Hz, $\text{H}^{1'}$), 4.43 (4H, m, H^{17}), 4.26 (1H, m, $\text{H}^{3'}$), 4.08-4.05 (3H, m, H^{18} plus $\text{H}^{4'}$), 3.66 (6H, s, OCH_3), 3.20 (2H, m, $\text{H}^{5'}$), 2.30 (1H, m, $\text{H}^{2'_{\alpha}}$), 1.91 (1H, m, $\text{H}^{2'_{\beta}}$), 1.66 (3H, s, CH_3).

^{13}C NMR (100 MHz, DMSO): δ 158.1, 150.6, 148.1, 145.0 (C^{Ar}), 144.2 (C^6), 143.7, 143.1, 140.7, 138.5 (C^{Ar}), 137.4 (C^4), 137.2, 135.7, 135.6 (C^{Ar}), 130.4, 129.7, 127.8, 127.7, 127.6, 126.9, 126.7, 126.7, 124.7, 120.0 (CH^{Ar}), 113.2 (C^{10}), 86.3 ($\text{C}^{4'}$), 85.5 (C^7), 76.9 ($\text{C}^{1'}$), 72.6 ($\text{C}^{3'}$), 67.9 (C^{17}), 64.6 ($\text{C}^{5'}$), 55.0, 55.0 (OCH_3), 45.9 (C^{18}), 43.6 ($\text{C}^{2'}$), 16.0 (CH_3).

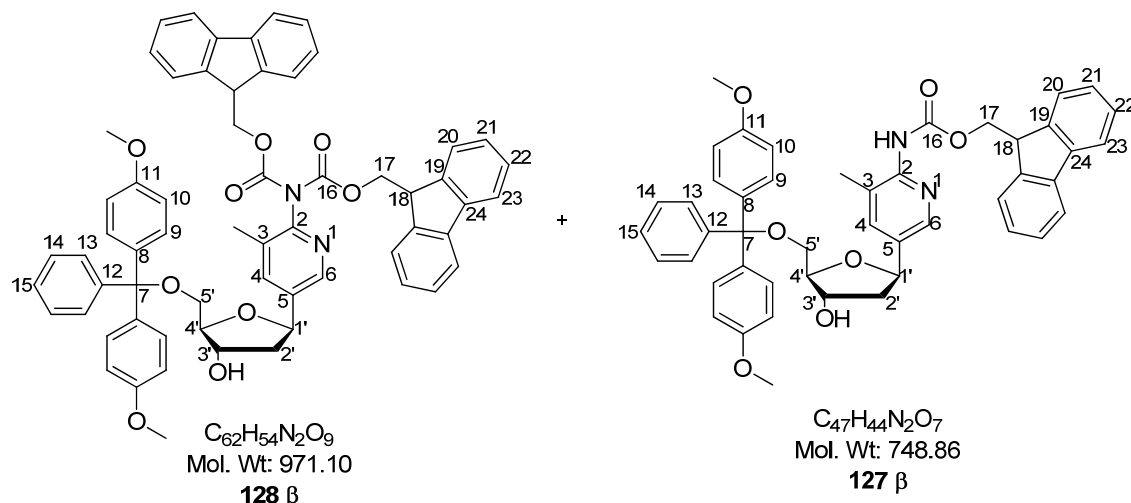
Data for 124 β :

R_f (Pet ether: EA 6:4, A'): 0.67

Data for 127 β :

R_f (Pet ether: EA 6:4, A'): 0.18

2-[*N,N*-bis-(fluorenylmethyloxycarbonyl)]amino-3-methyl-5-(5'-*O*-(4,4'-dimethoxytrityl)-2'-deoxy- β -D-ribofuranosyl)pyridine, (**128** β) and 2-[*N*-(fluorenylmethyloxycarbonyl)]amino-3-methyl-5-(5'-*O*-(4,4'-dimethoxytrityl)-2'-deoxy- β -D-ribofuranosyl)pyridine, (**127** β)



To a solution of **125** β (0.0495 g, 0.0473 mmol) in anhydrous pyridine (0.5 mL) was added a solution of KF (aq, 1.0 M, 0.1 mL, 0.0946 mmol) and the reaction mixture left to stir at r.t for 1 hr 30 mins after which reaction mixture was partitioned between DCM and water. The organic phases were combined and dried over anhydrous sodium sulphate. Solvent was removed *in vacuo*. Column chromatography (Hex: DCM (0.5 % Pyridine) 2:8, 0:1, DCM: EA 9:1, 8:2, 7:3, 1:1, 0:1 v/v) afforded **128** β (0.022 g, 0.023 mmol) and **127** β (0.0094 g, 0.013 mmol) as white foams (76 % overall yield).

Data for **128** β :

R_f (DCM: Hex 8:2, A'): 0.30

Data for **127** β :

R_f (DCM: Hex 8:2, A'): 0.05

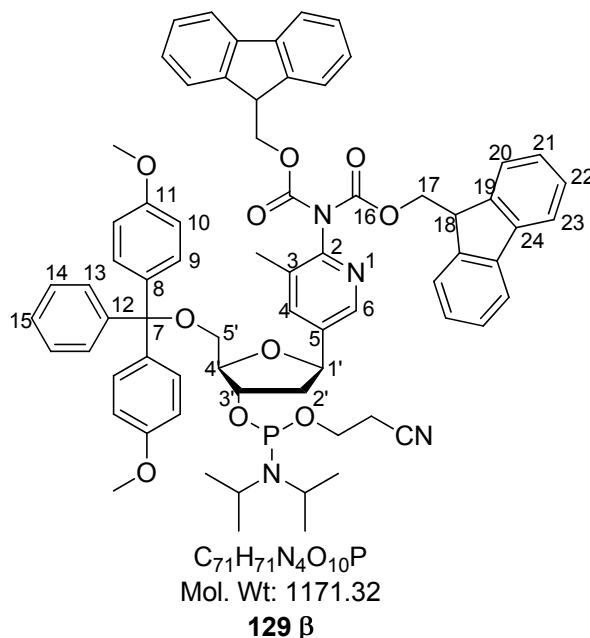
LRMS: $[ES^+, MeOH]$ m/z (%): 772 ($(M(^{13}C) + Na)^+$, 45), 771 ($(M + Na)^+$, 90), 750 ($(M(^{13}C) + H)^+$, 12), 749 ($(M + H)^+$, 100).

HRMS: [ES⁺] Found 749.3220 Da (M+H)⁺, calculated 749.3221 Da.

¹H NMR (400 MHz, DMSO): δ 9.58 (1H, s, NH), 8.26 (1H, d, *J* = 2.0 Hz, H⁶), 7.88 (2H, d, *J* = 7.5 Hz, H^{Ar}), 7.71 (2H, d, *J* = 6.5 Hz, H^{Ar}), 7.66 (1H, d, *J* = 1.5 Hz, H⁴), 7.43-7.14 (13H, m, H^{Ar}), 6.87 (4H, d, *J* = 8.6 Hz, H¹⁰), 5.17 (1H, d, *J* = 3.5 Hz, OH^{3'}), 5.11 (1H, dd, *J* = 10.0, 5.5 Hz, H^{1'}), 4.39 (4H, d, *J* = 7.0 Hz, H¹⁷), 4.27 (1H, t, *J* = 7.0 Hz, H¹⁸), 4.20 (1H, m, H^{3'}), 3.96 (1H, m, H^{4'}), 3.72 (6H, s, OCH₃), 3.15 (1H, dd, *J* = 10.5, 4.0 Hz, H^{5'}), 3.11 (1H, dd, *J* = 10.5, 5.0 Hz, H^{5'}), 2.18 (1H, m, H^{2'a}), 2.09 (3H, s, CH₃), 1.93 (1H, m, H^{2'p}).

¹³C NMR (100 MHz, DMSO): δ 158.1 (C¹⁶), 154.0 (C¹¹), 149.0 (C²), 144.9 (C⁶), 143.7, 143.6 (Fmoc-C^{Ar}), 140.7 (DMT-C^{Ar}), 137.0 (C⁴), 135.7 (C⁵), 129.7 (CH^{Ar}), 128.1 (DMT-C^{Ar}), 127.8 (C³), 127.7, 127.7, 127.0, 126.7, 120.0 (CH^{Ar}), 113.2 (C¹⁰), 86.3 (C^{4'}), 85.4 (C⁷), 76.9 (C^{1'}), 72.6 (C^{3'}), 65.9 (C¹⁷), 64.5 (C^{5'}), 55.0 (OCH₃), 46.6 (C¹⁸), 43.7 (C^{2'}), 17.3 (CH₃).

2-[*N,N*-bis-(fluorenylmethyloxycarbonyl)]amino-3-methyl-5-(5'-*O*-(4,4'-dimethoxytrityl)-2'-deoxy- β -D-ribofuranosyl)pyridine-3'-*O*-(2-cyanoethyl-*N,N*-diisopropyl)phosphoramidite, (129 β**)²⁷⁵**



To a solution of **128 β** (0.1637 g, 0.169 mmol) in anhydrous DCM (0.5 mL) was added *N,N*-diisopropylethylamine (0.06 mL, 0.337 mmol) followed by dropwise addition of 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite (0.37 M in DCM, 0.5 mL, 0.185 mmol) under an argon atmosphere, and the reaction mixture left to stir at r.t for 1 hr 55 mins, after which another 0.2 eq (0.0075 mL, 0.034 mmol) phosphitilating reagent was added and reaction mixture was left to stir at r.t for another 35 mins. The reaction mixture was then transferred *via* needle and syringe to a separating funnel containing anhydrous DCM. The reaction mixture was washed with saturated KCl (5 mL) and the organic phase was dried over anhydrous sodium sulphate. Solvent was removed *in vacuo* under argon, to give following purification by column chromatography (Hex: EA 7:3 (0.5 % pyridine)) under argon pressure, **129 β** , a diastereoisomeric mixture (1:1) as a white foam (0.0674 g, 34 %).

Data for the diastereoisomeric mixture:

R_f (Hex: EA 1:1 (0.5 % pyridine), A'): 0.47, 0.28.

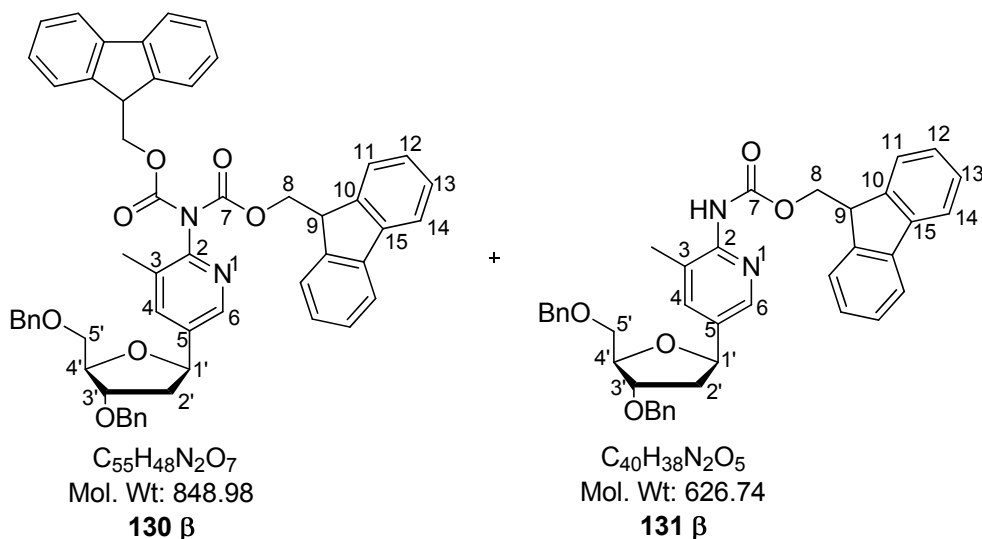
LRMS: [ES⁺, MeOH] *m/z* (%): 1196 ((M(2 ¹³C) + Na)⁺, 25), 1195 ((M(¹³C) + Na)⁺, 78), 1194 ((M + Na)⁺, 100), 1172 ((M + H)⁺, 18).

³¹P NMR (121 MHz, CD₃CN) δ 148.90 (s, 1P, **P**), 148.75 (s, 1P, **P**).

¹H NMR (400 MHz, CD₃CN): δ 8.04 (2H, s, **H**⁶), 7.66 (8H, d, *J* = 8.0 Hz, **H**^{Ar}), 7.51-7.48 (4H, m, **H**^{Ar}), 7.45 (2H, m, **H**⁴), 7.37-7.19 (38H, m, **H**^{Ar}), 6.82-6.79 (8H, m, **H**¹⁰), 5.19 (2H, dd, *J* = 8.0, 4.0 Hz, **H**^{1'}), 4.57 (2H, m, **H**^{3'}), 4.46 (8H, m, **H**¹⁷), 4.29-4.26 (2H, m, **H**^{4'}), 4.03 (4H, t, *J* = 8.0 Hz, **H**¹⁸), 3.91-3.78 (2H, m, OCH₂CH₂CN), 3.76-3.73 (2H, m, OCH₂CH₂CN), 3.69-3.59 (16H, s, OCH₃ plus iPr-CH), 3.32 (2H, t, *J* = 4.0 Hz, **H**^{5'}), 3.28 (2H, d, *J* = 4.0 Hz, **H**^{5'}), 2.68 (2H, t, *J* = 8.0 Hz, OCH₂CH₂CN), 2.58 (2H, t, *J* = 8.0 Hz, OCH₂CH₂CN), 2.53-2.52 (1H, m, **H**^{2'a}), 2.45 (1H, dd, *J* = 16.0, 8.0 Hz, **H**^{2'a}), 2.09-2.01 (2H, m, **H**^{2'β}), 1.55 (6H, s, CH₃), 1.22-1.13 (24H, m, iPr-CH₃).

¹³C NMR (100 MHz, CD₃CN): δ 159.4 (**C**¹¹), 151.7 (**C**²), 149.4 (**C**^{Ar}), 146.0 (**C**⁶), 145.1, 144.2, 141.8 (Fmoc-**C**^{Ar}), 138.7, 138.1 (**C**⁴), 136.8, 136.8 (DMT-**C**^{Ar}), 131.7, 130.8, 128.8, 128.6, 128.4, 127.8, 127.6, 125.4, 120.8 (CH^{Ar}), 113.8 (**C**¹⁰), 86.9 (**C**^{4'}), 78.4, 78.3 (**C**^{1'}), 76.3, 76.2 (**C**^{3'}), 68.5 (**C**¹⁷), 65.1, 65.1 (**C**^{5'}), 59.4, 59.2 (OCH₂CH₂CN), 55.6 (OCH₃), 47.1 (**C**¹⁸), 43.9 (iPr-CH), 43.8, 43.5 (**C**^{2'}), 24.8, 24.7 (iPr-CH₃), 20.9 (OCH₂CH₂CN), 16.5 (CH₃).

2-[*N,N*-bis-(fluorenylmethyloxycarbonyl)]amino-3-methyl-5-(3',5'-di-*O*-benzyl-2'-deoxy- β -D-ribofuranosyl)pyridine, (**130** β) and 2-[*N*-(fluorenylmethyloxycarbonyl)]amino-3-methyl-5-(3',5'-di-*O*-benzyl-2'-deoxy- β -D-ribofuranosyl)pyridine, (**131** β)



A solution of **107 β** (1.0777 g, 2.66 mmol) in anhydrous pyridine (7.0 mL) was cooled to 0 °C before careful addition of Fmoc-Cl (5.514 g, 21.3 mmol) in anhydrous acetonitrile (5.0 mL). The reaction mixture was left to stir at 0 °C for 2 hrs (reaction mixture had turned dark brown from mustard yellow) after which solvent was reduced under high vacuum. The residue was then partitioned between water and DCM. The organic phases were combined and dried over anhydrous sodium sulphate. Solvent was removed *in vacuo* and high vacuum to give crude (5.8508 g) as a mustard yellow oil. Column chromatography (Hex: EA 9:1, 8:2, 7:3, 6:4, 1:1, 0:1, EA: MeOH 8:2 v/v) afforded **130 β** (0.6123 g, 0.722 mmol), and a mixture of **130 β** and **131 β** (3:7 ratio respectively, 1.26 g), as cream white foams. The mixture was put for further reaction to convert **130 β** into **131 β** using this same method, and then purification by column chromatography (Hex: EA 9:1, 8:2, 7:3, 6:4, 1:1, 0:1, EA: MeOH 8:2 v/v) afforded **130 β** (0.5448 g, 0.642 mmol) and **131 β** (0.7433 g, 1.186 mmol) as cream white foams. Overall yield for this reaction was 95.9 %.

Data for 130 β :

R_f (Hex: EA 1:1; A'): 0.39.

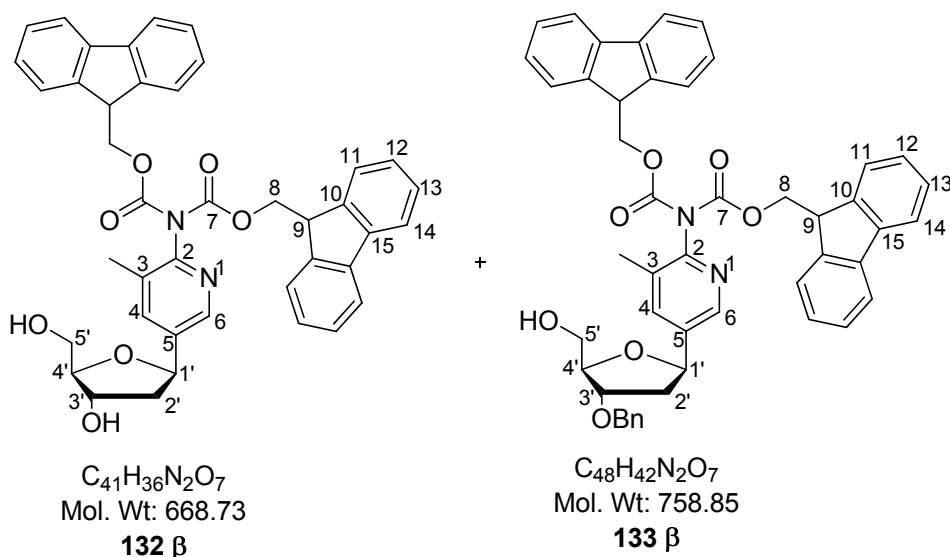
LRMS: [ES⁺, MeOH] *m/z* (%): 850 ((M(¹³C) + H)⁺, 44), 849 ((M + H)⁺, 100).

HRMS: [ES⁺] Found 871.3367 Da (M + Na)⁺, calculated 871.3354 Da.

¹H NMR (400 MHz, CDCl₃): δ 8.21 (1H, d, *J* = 2.0 Hz, **H⁶**), 7.66 (4H, d, *J* = 7.6 Hz, **H^{Ar}**), 7.56 (1H, d, *J* = 2.0 Hz, **H⁴**), 7.43-7.28 (17H, m, **H^{Ar}**), 7.20 (5H, m, **H^{Ar}**), 5.23 (1H, dd, *J* = 10.5, 5.0 Hz, **H^{1'}**), 4.64 (2H, s, Bn-CH₂), 4.63 (2H, s, Bn-CH₂), 4.49-4.40 (5H, m, **H⁸** plus **H^{4'}**), 4.25 (1H, d, *J* = 5.5 Hz, **H^{3'}**), 4.09-4.06 (2H, m, **H⁹**), 3.73 (1H, dd, *J* = 10.0, 4.0 Hz, **H^{5'}**), 3.67 (1H, dd, *J* = 10.0, 5.0 Hz, **H^{5'}**), 2.45 (1H, dd, *J* = 12.0, 5.5 Hz, **H^{2'}_w**), 1.96 (1H, ddd, *J* = 13.5, 11.0, 6.0 Hz, **H^{2'}_p**), 1.83 (3H, s, CH₃).

¹³C NMR (100 MHz, CDCl₃): δ 151.6 (**C^{Ar}**), 148.9 (**C²**), 144.9 (**C⁶**), 144.5, 143.4 (Fmoc-**C^{Ar}**), 141.7, 141.3, 138.4 (**C^{Ar}**), 138.2, 138.1 (Bn-**C^{Ar}**), 137.6 (**C⁴**), 131.3 (**C³**), 128.7, 128.6, 128.0, 127.9, 127.8, 127.8, 127.2, 127.2, 120.2, 120.1 (CH^{Ar}), 84.3 (**C^{4'}**), 81.7 (**C^{3'}**), 77.9 (**C^{1'}**), 73.8, 71.4 (Bn-CH₂), 71.3 (**C^{5'}**), 68.8 (**C⁸**), 46.7 (**C⁹**), 41.5 (**C^{2'}**), 16.8 (CH₃).

2-[*N,N*-bis-(fluorenylmethyloxycarbonyl)]amino-3-methyl-5-(2'-deoxy- β -D-ribofuranosyl)pyridine, (**132** β) and 2-[*N,N*-bis-(fluorenylmethyloxycarbonyl)]amino-3-methyl-5-(3'-*O*-benzyl-2'-deoxy- β -D-ribofuranosyl)pyridine, (**133** β)²⁷⁷



A solution of **130 β** (1.0382 g, 1.22 mmol) in anhydrous DCM (3.7 mL) was cooled to - 78 °C before careful dropwise addition of boron trichloride (1.0 M in DCM, 3.7 mL, 3.7 mmol). The reaction mixture changed colour from mustard yellow to dark brown. The reaction mixture was left to stir at - 78 °C for 6 hrs after which methanol (HPLC grade, ~ 100 mL) was added and the reaction mixture was left to stir at - 75 °C for 30 mins after which it was left to stand at 4 °C for 17 hrs 30 mins. Sat. $NaHCO_3$ was added until pH 7.0 obtained. Solvent was removed *in vacuo* and the residue partitioned between DCM and water. The organic phases were combined and dried over anhydrous Na_2SO_4 . Solvent was removed *in vacuo* to give crude as a cream foam. Column chromatography (DCM: MeOH 1:0, 98:2, 96:4, 94:6, 92:8, 9:1 v/v) afforded **132 β** (0.542 g) and **133 β** (0.1162 g) as cream foams (overall yield 78.6 %).

Data for **132 β** :

R_f (DCM: MeOH 9:1; A'): 0.43.

LRMS: [ES⁺, MeOH] *m/z* (%): 692 ((M(¹³C) + Na)⁺, 42), 691 ((M + Na)⁺, 100), 669 ((M + H)⁺, 12).

HRMS: [ES⁺] Found 691.2407 Da (M + Na)⁺, calculated 691.2415 Da.

¹H NMR (400 MHz, DMSO): δ 8.17 (1H, s, **H**⁶), 7.77 (3H, d, *J* = 7.0 Hz, Fmoc-**H**^{Ar}), 7.62 (1H, s, **H**⁴), 7.35 (4H, m, Fmoc-**H**^{Ar}), 7.28-7.17 (9H, m, Fmoc-**H**^{Ar}), 5.18-5.13 (2H, m, **OH**^{3'} plus **H**^{1'}), 4.85 (1H, t, *J* = 5.5 Hz, **OH**^{5'}), 4.44 (4H, m, **H**⁸), 4.31 (1H, m, **H**^{3'}), 4.10-4.07 (2H, m, **H**⁹), 3.92-3.91 (1H, m, **H**^{4'}), 3.61-3.51 (2H, m, **H**^{5'}), 2.23 (1H, dd, *J* = 11.5, 5.0 Hz, **H**^{2'}_a), 1.89 (1H, td, *J* = 12.6, 5.5 Hz, **H**^{2'}_β), 1.73 (3H, s, **CH**₃).

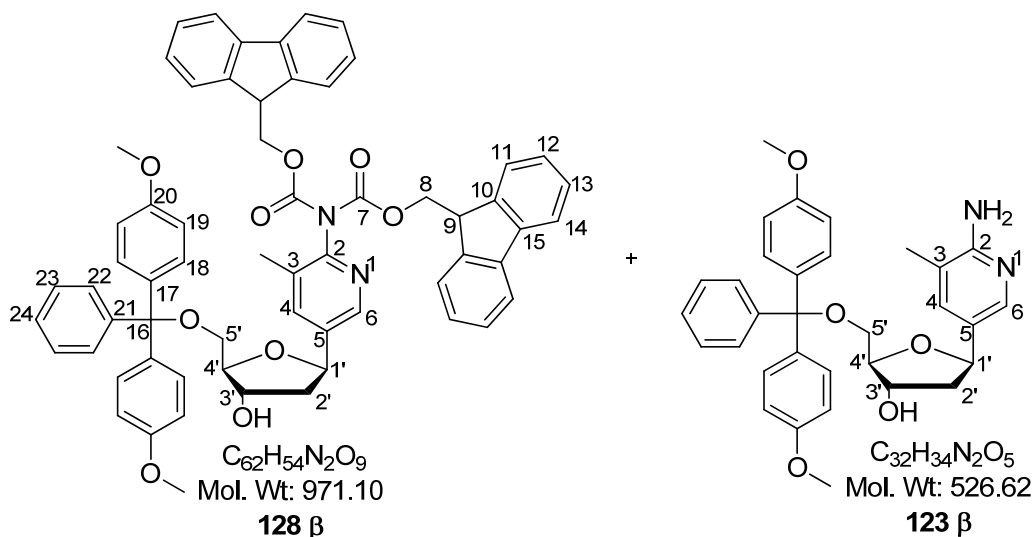
¹³C NMR (100 MHz, DMSO): δ 150.7 (**C**^{Ar}), 148.0 (**C**²), 144.3 (**C**⁶), 143.7, 143.1, 140.6 (Fmoc-**C**^{Ar}), 138.8 (**C**⁵), 137.4 (**C**⁴), 130.4 (**C**³), 128.9, 128.6, 128.4, 128.2, 127.6, 127.3 (Fmoc-**CH**^{Ar}), 126.9 (Fmoc-**C**^{Ar}), 125.9, 125.2 (Fmoc-**CH**^{Ar}), 124.7 (Fmoc-**C**^{Ar}), 121.4, 120.0 (Fmoc-**CH**^{Ar}), 88.1 (**C**^{4'}), 76.8 (**C**^{1'}), 72.5 (**C**^{3'}), 68.0 (**C**⁸), 62.4 (**C**^{5'}), 45.9 (**C**⁹), 43.4 (**C**^{2'}), 16.0 (**CH**₃).

Data for 133 β:

R_f (DCM: MeOH 9:1; A[']): 0.69.

LRMS: [ES⁺, MeOH] *m/z* (%): 781 ((M + Na)⁺, 100).

2-[*N,N*-bis-(fluorenylmethoxycarbonyl)]amino-3-methyl-5-(5'-*O*-(4,4'-dimethoxytrityl)-2'-deoxy- β -D-ribofuranosyl)pyridine, (**128 β**) and 2-amino-3-methyl-5-(5'-*O*-(4,4'-dimethoxytrityl)-2'-deoxy- β -D-ribofuranosyl)pyridine, (**123 β**)¹⁵⁸



Co-evaporation of **132 β** (0.70 g, 1.048 mmol) with anhydrous pyridine (3 x 3 mL) was carried out under high vacuum and then left to dry overnight under high vacuum. To a solution of **132 β** (0.70 g, 1.048 mmol) in anhydrous pyridine (2.5 mL) was added a solution of 4,4'-dimethoxytrityl chloride (0.40 g, 1.18 mmol) in anhydrous pyridine (2.5 mL) dropwise and the reaction mixture was left to stir at r.t for 2 hrs after which 0.1 eq 4,4'-dimethoxytrityl chloride (0.035 g, 0.105 mmol) in anhydrous pyridine (0.3 mL) was added and the reaction mixture left to stir at r.t for 1 hr. Another 0.15 eq 4,4'-dimethoxytrityl chloride (0.053 g, 0.156 mmol) in anhydrous pyridine (0.3 mL) was added and the reaction mixture left to stir at r.t for 30 mins. Solvent was reduced to half under high vacuum, methanol (2 mL) and TEA (1 mL) were added, and the reaction mixture partitioned between DCM and a 1:1 mixture of H₂O: sat. NaHCO₃ aq. The organic phases were combined and dried over anhydrous sodium sulphate. Solvent was removed *in vacuo* and high vacuum to give crude as a pale orange foam. Column chromatography (Hex: EA (0.5 % Pyridine) 1:1, 4:6, EA: MeOH 9:1 v/v) afforded **128 β** (0.2211 g, 0.230 mmol) and **123 β** (0.36 g, 0.684 mmol) as white foams (87.2 % overall yield).

Data for 128 β :

R_f (Hex: EA 1:1(0.5 % Pyridine); A[']): 0.15.

LRMS: [ES⁺, MeOH] *m/z* (%): 994 ((M(¹³C) + Na)⁺, 62), 993 ((M + Na)⁺, 100), 973 ((M(¹³C) + H)⁺, 3), 972 ((M + H)⁺, 5).

HRMS: [ES⁺] Found 971.3898 Da (M + H)⁺, calculated 971.3902 Da.

¹H NMR (400 MHz, DMSO): δ 8.15 (1H, d, *J* = 2.0 Hz, **H⁶**), 7.74 (4H, d, *J* = 7.5 Hz, **H^{Ar}**), 7.57 (1H, d, *J* = 1.5 Hz, **H⁴**), 7.45 (4H, d, *J* = 7.0 Hz, **H^{Ar}**), 7.32-7.19 (19H, m, **H¹⁸** plus **H^{Ar}**), 6.83 (4H, d, *J* = 8.5 Hz, **H¹⁹**), 5.26 (1H, d, *J* = 4.5 Hz, **OH^{3'}**), 5.20 (1H, dd, *J* = 10.0, 5.5 Hz, **H^{1'}**), 4.44 (4H, m, **H⁸**), 4.26 (1H, m, **H^{3'}**), 4.08-4.05 (3H, m, **H⁹** plus **H^{4'}**), 3.67 (6H, s, **OCH₃**), 3.20 (2H, d, *J* = 4.5 Hz, **H^{5'}**), 2.31-2.27 (1H, m, **H^{2'a}**), 2.00-1.90 (1H, m, **H^{2'β}**), 1.66 (3H, s, **CH₃**).

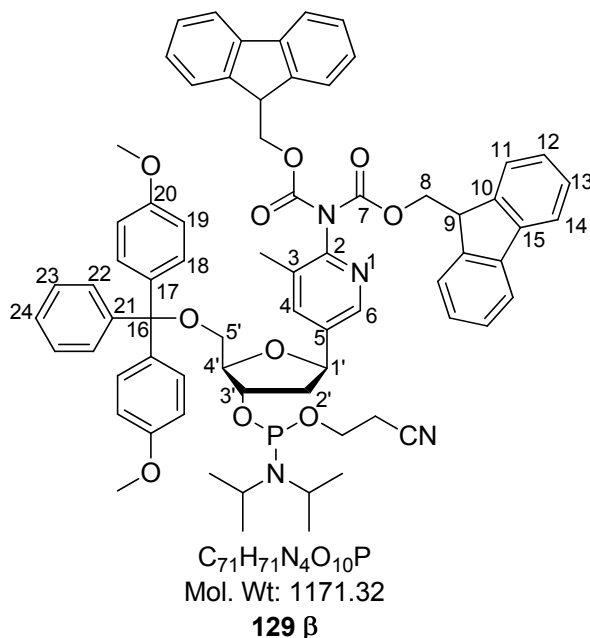
¹³C NMR (100 MHz, DMSO): δ 158.0 (**C²⁰**), 150.6 (DMT-**C^{Ar}**), 148.1 (**C²**), 144.9 (DMT-**C^{Ar}**), 144.2 (**C⁶**), 143.1, 140.6 (Fmoc-**C^{Ar}**), 138.5 (**C^{Ar}**), 137.2 (**C⁴**), 135.7, 135.6 (**C^{Ar}**), 130.4 (**C³**), 129.7, 127.8, 127.7, 127.6, 126.9, 126.7 (**CH^{Ar}**), 124.7 (Fmoc-**CH^{Ar}**), 120.0 (**CH^{Ar}**), 113.2 (**C¹⁹**), 86.3 (**C^{4'}**), 85.5 (**C¹⁶**), 76.9 (**C^{1'}**), 72.5 (**C^{3'}**), 67.9 (**C⁸**), 64.6 (**C^{5'}**), 54.9, 54.9 (**OCH₃**), 45.9 (**C⁹**), 43.6 (**C^{2'}**), 16.0 (**CH₃**).

Data for 123 β :

R_f (Hex: EA 1:1(0.5 % Pyridine); A[']): 0.08.

All other analytical data are consistent with that reported previously for this compound.

2-[*N,N*-bis-(fluorenylmethyloxycarbonyl)]amino-3-methyl-5-(5'-*O*-(4,4'-dimethoxytrityl)-2'-deoxy- β -D-ribofuranosyl)pyridine-3'-*O*-(2-cyanoethyl-*N,N*-diisopropyl)phosphoramidite, (129 β**)²⁷⁵**



To a solution of **128 β** (0.387 g, 0.3985 mmol) in anhydrous DCM (2.3 mL) was added anhydrous *N,N*-diisopropylethylamine (0.14 mL, 0.797 mmol) followed by dropwise addition of 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite (0.1 mL, 0.4383 mmol) under an argon atmosphere, and the reaction mixture left to stir at r.t for 1 hr 25 mins after which another 0.11 eq (0.01 mL, 0.0448 mmol) phosphitilating reagent was added and reaction mixture was left to stir at r.t for another 35 mins. The reaction mixture was transferred *via* needle and syringe to a separating funnel containing anhydrous DCM. The reaction mixture was washed with saturated KCl (10 mL) and the organic phase was dried over anhydrous sodium sulphate. Solvent was removed *in vacuo* under argon, to give following purification by column chromatography (Hex: EA 1:1 (0.5 % pyridine) v/v) under argon pressure, **129 β** , as a white foam (0.231 g, 49.5 %, diastereoisomers 1:1).

Data for the diastereoisomeric mixture:

R_f (Hex: EA 1:1 (0.5 % pyridine), A'): 0.27; UV positive; Stains red-orange.

LRMS: [ES⁺, MeOH] *m/z* (%): 1194 ((M(¹³C) + Na)⁺, 30), 1193 ((M + Na)⁺, 48), 1174 ((M (3 ¹³C) + H)⁺, 5), 1173 ((M (2 ¹³C) + H)⁺, 18), 1172 ((M (¹³C) + H)⁺, 65), 1171 ((M + H)⁺, 100).

³¹P NMR (121 MHz, CD₃CN) δ 148.90 (s, 1P, **P**), 148.75 (s, 1P, **P**).

¹H NMR (400 MHz, CD₃CN): δ 8.04 (2H, s, **H**⁶), 7.66 (8H, d, *J* = 7.5 Hz, **H**^{Ar}), 7.51-7.45 (6H, m, **H**^{Ar} plus **H**⁴), 7.38-7.19 (38H, m, **H**^{Ar} plus **H**¹⁸), 6.83-6.80 (8H, m, **H**¹⁹), 5.20 (2H, dd, *J* = 9.0, 4.0 Hz, **H**^{1'}), 4.59-4.56 (2H, m, **H**^{3'}), 4.47 (8H, m, **H**⁸), 4.29-4.26 (2H, m, **H**^{4'}), 4.05-4.02 (4H, m, **H**⁹), 3.87-3.84 (2H, m, OCH₂CH₂CN), 3.75-3.74 (2H, m, OCH₂CH₂CN), 3.69 (16H, m, OCH₃ plus iPr-CH), 3.33 (2H, t, *J* = 4.5 Hz, **H**^{5'}), 3.30 (2H, d, *J* = 4.5 Hz, **H**^{5'}), 2.68 (2H, t, *J* = 6.0 Hz, OCH₂CH₂CN), 2.58 (2H, t, *J* = 6.0 Hz, OCH₂CH₂CN), 2.55-2.43 (2H, m, **H**^{2'}_α), 2.10-2.04 (2H, m, **H**^{2'}_β), 1.58 (6H, s, CH₃), 1.23-1.13 (24H, m, iPr-CH₃).

¹³C NMR (100 MHz, CD₃CN): δ 159.7 (**C**²⁰), 151.9 (**C**²), 149.6, 146.2 (**C**^{Ar}), 145.3 (**C**⁶), 144.4 (Fmoc-**C**^{Ar}), 142.1, 139.0, 139.0 (**C**^{Ar}), 138.3 (**C**⁴), 131.9 (**C**³), 131.1, 129.1, 129.0, 128.9, 128.6, 128.1, 128.0, 126.9, 125.6, 124.8, 121.0 (CH^{Ar}), 114.0 (**C**¹⁹), 87.1, 86.7 (**C**^{4'}), 78.6, 78.5 (**C**^{1'}), 77.0, 76.4 (**C**^{3'}), 68.8 (**C**⁸), 65.4, 65.3 (**C**^{5'}), 59.5 (OCH₂CH₂CN), 55.9 (OCH₃), 47.3 (**C**⁹), 44.1 (iPr-CH), 44.0, 43.7 (**C**^{2'}), 24.9 (iPr-CH₃), 21.1 (OCH₂CH₂CN), 16.7 (CH₃).

5.3 OLIGONUCLEOTIDE SYNTHESIS

All phosphoramidites synthesised were incorporated into various oligonucleotides sequences and then analysed by negative mode electrospray MS analysis.

5.3.1 SYNTHESIS OF OLIGONUCLEOTIDES

The synthesised monomers were treated as follows: After purification by column chromatography the monomer was dissolved in freshly distilled DCM. Aliquots corresponding to between ~60-100 μ moles (depending on the phosphitilation yields) were transferred to ABI monomer reagent bottles and dried in a desiccator overnight under high vacuum before being stored under a positive pressure of argon at -20°C .

Standard DNA phosphoramidites, solid supports, and additional reagents were purchased from Link Technologies or Applied Biosystems, Ltd. The psoralen C6 phosphoramidite was purchased from Glen Research, Inc. All oligonucleotides were synthesized on an Applied Biosystems 394 automated DNA/RNA synthesizer using a 1.0 μ mol phosphoramidite cycle of acid-catalyzed detritylation, coupling, capping, and iodine oxidation. Normal monomers (A, G, C, and T) were allowed to couple for 25 seconds and all other monomers for an additional 300 seconds. Stepwise coupling efficiencies and overall yields of monomers with DMT protection were determined by measuring trityl cation conductivity and in all cases were $> 98.0\%$. All β -cyanoethyl phosphoramidite monomers were dissolved in anhydrous acetonitrile to a concentration of 0.1 M immediately prior to use. To all psoralen C6 containing oligonucleotides, the oligo was synthesised first, after which psoralen was incorporated as a phosphoramidite and deprotection followed.

The oligonucleotides attached to the synthesis columns were treated with 20 % diethylamine in acetonitrile for 20 mins before cleavage from the support. This procedure removes cyanoethyl groups from the phosphotriesters and scavenges the resultant acrylonitrile; preventing cyanoethyl adducts being formed at the primary amine sites of the modified nucleosides.

Cleavage of the oligonucleotides containing modified nucleosides, from the solid support and deprotection was achieved by exposure to concentrated aq. ammonia for 24 hrs at r.t (1 hr at r.t cleaved the oligos from the resin, the ammonia solution was then left at r.t for an additional 23 hrs to allow for deprotection). Cleavage of the oligonucleotides containing non-modified nucleosides, from the solid support and deprotection was achieved quantitatively by treatment with concentrated aqueous ammonia at 55 °C for 5 hrs.

5.3.2 OLIGONUCLEOTIDE PURIFICATION

Purification of oligonucleotides was carried out by reversed-phase HPLC by using a Gilson system with a Brownlee Aquapore column (C8, 8 mm x 250 mm, 300 Å pore) with a gradient of CH₃CN in NH₄OAc increasing from 0 to 50 % buffer B over 30 mins with a flow rate of 4 mLmin⁻¹ (elution buffer A: 0.1 M NH₄OAc, pH 7.0; buffer B: 0.1 M NH₄OAc with 50 % CH₃CN, pH 7.0). Elution of oligonucleotides was monitored by ultraviolet absorption at 295 nm. After HPLC purification, oligonucleotides were desalted by using NAP-10 Sephadex columns (GE Healthcare) according to the manufacturer's instructions.

The eluted oligonucleotides were analysed by UV-Visible absorption spectroscopy (295 nm) and the concentrations of each oligonucleotide were calculated from the standard molar extinction coefficients of the four natural bases ($\epsilon = \text{mol}^{-1} \text{ L cm}^{-1}$ for the oligonucleotide) = (A 15.4) + (G 11.5) + (C 7.4) + (T 8.7) $\times 0.91$)³⁰⁵. The oligonucleotides were then stored under argon permanently at - 20 °C or temporarily at 5 °C.

5.3.3 OLIGONUCLEOTIDE ANALYSIS

Oligonucleotide analysis was carried out by capillary gel electrophoresis (CE) and mass spectrometry. Each sample (0.4 OD per 100 µL) was injected on to an ssDNA 100-R gel by using a Tris-borate urea (7 M) buffer (kit N° 477480) on a Beckman coulter P/ACETM MDQ capillary gel electrophoresis system controlled by 32 Karat software. The injection voltage was 10.0 kV and separation voltage was 9.0 kV. The

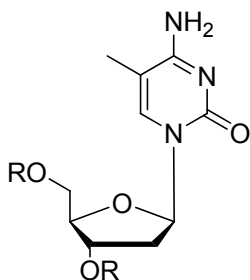
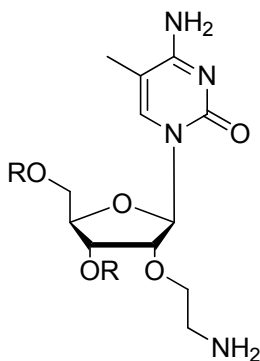
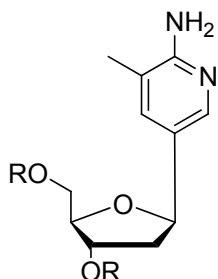
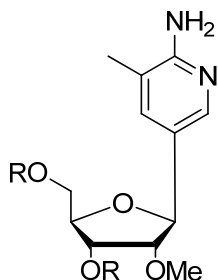
5.3.4.2 DUPLEXES

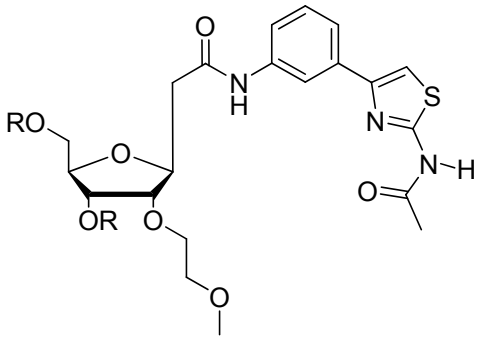
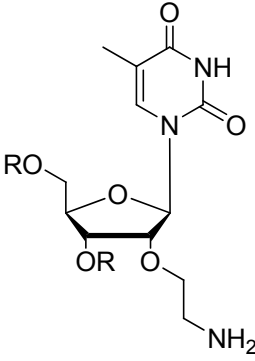
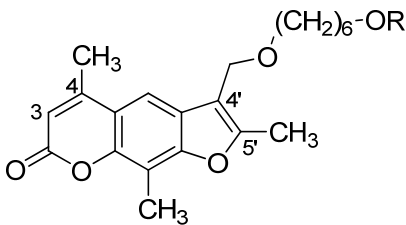
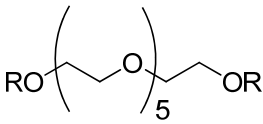
Oligo code./ Serial No.	Sequence d(5'-3')	Molecular ion expected (M - H) ⁻	Molecular ion found (M - H) ⁻
OL7	ACTTAGGTGAAGAAGAAGCAGAAGAAGAAG	9418.0	9418.0
OL8	CTTCTTCTTCTGCTTCTTCTTCACCTAAGT	8990.0	8991.0

5.3.4.3 HAIRPIN DUPLEXES

Oligo code./ Serial No.	Sequence d(5'-3')	Molecular ion expected (M + H) ⁺	Molecular ion found (M + H) ⁺
OL9	GACAAAAAGCGAAAAAGTAGHCTACTTTTTCCGCTTTTTGTC	13256.0	13260.9

5.3.4.4 KEY TO MODIFIED NUCLEOSIDES AND OLIGONUCLEOTIDE MODIFICATIONS

Code	Name	Structure
^{Me} C	5-methyl deoxy Cytosine	
^{Me} C _{AE}	5-methyl-(2'-O-aminoethyl-β-D-ribofuranosyl)cytosine	
^{Me} P	3-methyl-2-amino-5-(2'-deoxy-β-D-ribofuranosyl)pyridine	
^{Me} P _{OMe}	3-methyl-2-amino-5-(2'-O-methyl-β-D-ribofuanosyl)pyridine	

S_{ME}	2'- <i>O</i> -methoxyethyl S	
T_{AE}	2'- <i>O</i> -aminoethyl thymidine	
Pso	Psoralen C6	
H	hexa(ethylene glycol)	

5.4 BIOPHYSICAL STUDIES

5.4.1 ULTRAVIOLET-MELTING ANALYSIS

UV melting experiments were performed on a Varian Cary 400 Scan UV-visible spectrophotometer, in Hellma SUPRASIL synthetic quartz, 10mm path length cuvettes, monitoring at 280 nm, with a DNA duplex concentration of 1.0 μ M and a volume of 1.2 mL. The sample chamber was flushed with air pumped through a desiccant to

prevent condensation of atmospheric water on the cells. Samples were prepared as follows: The third strand and duplex were mixed in a 3:1 ratio respectively in 2 mL Eppendorf tubes, then lyophilized before resuspending in 1.2 mL of the appropriate buffer solution (10 mM sodium phosphate at pH 6.2, 6.6, 7.0, 7.5 and 8.0, containing 200 mM NaCl). The samples were then filtered into the cuvettes with Kinesis regenerated cellulose 17 mm, 0.45 μ M syringe filters. The UV melting protocol (carried out pre-crosslinking for pH 6.2-8.0; post crosslinking for pH 7.0-8.0 only) involved initial denaturation by heating to 80 °C at 10 °C/min followed by annealing by cooling to 20 °C at 0.5 °C/min, then maintaining at 20 °C for 20 min before starting the melting experiment which involved heating from 20 °C to 80 °C at 0.5 °C/min, holding at 80 °C for two mins then cooling to 20 °C at 0.5 °C/min. Two successive melting curves were measured before fast annealing from 80 °C to 20 °C at 10 °C/min. Melting temperatures (T_m) were determined by the maximum of the first derivative plot of the second heating curve using OriginLab software.

5.4.1.1 BUFFER PREPARATIONS

Buffers at various pH were made according to the methods reported.³⁰⁷

5.4.1.1.1 Buffer pH 5.0

To a solution of acetic acid (0.12 g) in deionised water (180 mL) was added NaCl (2.238 g). The solution was then titrated to pH 5.5 at 20 °C with monovalent NaOH (0.1 M) and then made up to 200 mL with deionised water.

5.4.1.1.2 Buffer pH 6.2

To a solution of NaH_2PO_4 (0.24 g) in deionised water (180 mL) was added NaCl (2.178 g). The solution was then titrated to pH 6.19 with NaOH (0.1 M) and then made up to 200 mL with deionised water.

5.4.1.1.3 Buffer pH 6.6

To a solution of NaH_2PO_4 (0.24 g) in deionised water (180 mL) was added NaCl (2.134 g). The solution was then titrated to pH 6.59 with NaOH (0.1 M) and then made up to 200 mL with deionised water.

5.4.1.1.4 Buffer pH 7.0

To a solution of NaH_2PO_4 (0.24 g) in deionised water (180 mL) was added NaCl (2.081 g). The solution was then titrated to pH 7.0 with NaOH (0.1 M) and then made up to 200 mL with deionised water.

5.4.1.1.5 Buffer pH 7.5

To a solution of NaH_2PO_4 (0.24 g) in deionised water (180 mL) was added NaCl (2.029 g). The solution was then titrated to pH 7.5 with NaOH (0.1 M) and then made up to 200 mL with deionised water.

5.4.1.1.6 Buffer pH 8.0

To a solution of NaH_2PO_4 (0.24 g) in deionised water (180 mL) was added NaCl (2.004 g). The solution was then titrated to pH 8.0 with NaOH (0.1 M) and then made up to 200 mL with deionised water.

5.4.1.2 CROSS-LINKING

Triplexes were prepared as stated above (pH 7.0, 7.5 and 8.0 only) and each cuvette was suspended into an ice bath and irradiated with UV light (UVGL-58 handheld UV lamp, 365 nm) at a distance of 4 cm for 20 mins. The UV-melting method described above was then carried out post cross-linking.

5.4.2 FOOTPRINTING EXPERIMENTS

Footprinting studies were performed by Professor Keith Fox, Dr Nuria Vergara (School of Biological Sciences, University of Southampton) and me (School of Chemistry, University of Southampton).

5.4.2.1 FOOTPRINTING¹⁹⁹

5.4.2.1.1 GA track marker

The GA track marker was prepared as follows; 1.5 μL ^{32}P 3'-end radiolabelled DNA was mixed with 20 μL sterile water and 5 μL DNase I stop solution (10 mM EDTA, 1 mM NaOH, 0.2 % bromophenol blue, 80 % formamide). The sample was then incubated at 100 °C for 45 mins with the microcentrifuge tube cap open to allow evaporation.

5.4.2.1.2 DNase I footprinting

The footprinting buffers were prepared (50 mM sodium acetate, pH 5.5; 10 mM Tris-HCl, pH 7.0; 10 mM Tris-HCl, 50 mM NaCl, pH 7.5). 15 μL of the TFO of interest (**OL1**, 128.77 μM ; **OL2**, 98.66 μM ; **OL3**, 104.86 μM ; **OL4**, 74.01 μM) was taken for each different pH footprint and dried in a speed vacuum for ~ 20 mins until the pellet was completely dry. The pellet was then resuspended in the appropriate buffer to give a final concentration of 15 μM ; ~ 129 μL for **OL1**; ~ 99 μL for **OL2**; ~ 105 μL for **OL3** and ~ 74 μL for **OL4** (NOTE: in the final reaction, when the TFO is mixed with the DNA, the TFO will be at a final concentration of 10 μM). From the 10 μM TFO stock solution, serial dilutions were prepared to give 3.0 μM , 1.0 μM , 0.3 μM , 0.1 μM , 0.03 μM etc. In new Eppendorf tubes, 1.5 μL of ^{32}P 3'-end radiolabelled DNA (dissolved in TE2 buffer; 10 mM Tris, 0.1 mM EDTA, pH 7.5) was mixed with 3.0 μL TFO (dissolved in the appropriate buffer). An extra tube was prepared with just 1.5 μL radiolabelled DNA duplex and 3 μL footprinting buffer (negative control). The tubes including the GA track were left overnight at room temperature. After equilibration of the ligand–DNA complexes the mixtures were then digested; by adding to each

Eppendorf tube 2 μ L DNase I enzyme solution (4 μ L DNase I enzyme was diluted in 1 mL DNase I buffer (20 mM NaCl, 2 mM MgCl₂, 2 mM MnCl₂), 4 μ L of this initial diluted solution was then diluted again in 1 mL DNase I buffer). DNase I was purchased from Sigma and stored at -20 °C at a stock concentration of 7200 U/mL in 0.15 M NaCl containing 1 mM MgCl₂. After 1 min the digestion reaction was stopped by adding 4 μ L DNase I stop solution (10 mM EDTA, 1 mM NaOH, 0.1 % bromophenol blue, 80 % formamide) to each Eppendorf tube. Before loading onto the gel the DNA was denatured by incubating at 100 °C for 3 min and quickly cooled on ice. 6 μ L of each sample was then loaded onto the gel (8 % denaturing polyacrylamide gels containing 8 M urea, prepared in 1 x TBE buffer). Gels (40 cm long, 0.3 mm thick) were run at 1500 V for approximately 2 hrs until the dye reached the bottom of the plates. The gel plates were then separated, the gel was fixed by immersing in 10 % (v/v) acetic acid, transferred to Whatmann 3MM paper and dried under vacuum at 80 °C. The dried gel was then exposed to a phosphorimager screen overnight before scanning with a Storm phosphorimager using ImageQuant software. Bands were assigned by comparison with Maxam-Gilbert lanes specific for G and A.

5.4.2.2 PHOTODUCT FORMATION AND GEL ELECTROPHORESIS²⁹⁶

Samples were prepared in the same way as for footprinting. Triplexes were prepared with the radiolabelled duplex and TFO, **OL6** (varying concentrations; 10, 3, 1, 0.3, 0.1, 0.03 μ M) in either 10 mM Tris-HCl, pH 7.0 or 10 mM Tris-HCl, 50 mM NaCl pH 7.5, and were incubated at 20 °C in the dark overnight. The samples were then placed in a microtitre plate on ice and irradiated with UV light (6W UVGL-58 Mineralight lamp, 365 nm) at a distance of 2 cm for 20 minutes. The reaction was stopped by adding 4 μ L DNase I stop solution (10 mM EDTA, 1 mM NaOH, 0.1 % bromophenol blue, 80 % formamide). Samples were then heated to 100 °C for 3 minutes before rapidly cooling on ice, and loading onto a low percentage denaturing polyacrylamide gel (5 %). The polyacrylamide gel was run at 1500 V for ~ 2 hrs and then fixed in 10 % (v/v) acetic acid. The gel was then transferred to Whatman 3MM paper and dried under vacuum at 80 °C for 1 hr. The dried gel was then exposed to a phosphorimager screen overnight before scanning with a Storm phosphorimager.

CHAPTER 6

References

-
- (1) Brown, T. A. *Genomes*; 2nd ed.; BIOS Scientific Publishers Ltd: Oxford, **2002**.
 - (2) Dahm, R. *Hum. Genet.* **2008**, *122*, 565-581.
 - (3) Blackburn, G. M.; Gait, M. J. *Nucleic Acids in Chemistry and Biology*; 2nd ed.; Oxford University Press, **1996**.
 - (4) Levene, P. A. *J. Am. Chem. Soc.* **1910**, *32*, 231-240.
 - (5) Levene, P. A.; Jacobs, W. A. *Ber. Dtsch. Chem. Ges* **1909**, *42*, 2102-2106.
 - (6) Levene, P. A.; Jacobs, W. A. *Ber. Dtsch. Chem. Ges* **1909**, *42*, 3247-3251.
 - (7) Sundaralingam, M. *Biopolymers* **1969**, *7*, 821-860.
 - (8) Altona, C.; Sundaralingam, M. *J. Am. Chem. Soc.* **1972**, *94*, 8205-8212.
 - (9) Levene, P. A. *J. Biol. Chem.* **1919**, *40*, 415-424.
 - (10) Vischer, E.; Chargaff, E. *J. Biol. Chem.* **1947**, *168*, 781-782.
 - (11) Vischer, E.; Chargaff, E. *J. Biol. Chem.* **1948**, *176*, 703-714.
 - (12) Chargaff, E. *Experientia: Cellular and Molecular Life Sciences* **1950**, *6*, 201-209.
 - (13) Chargaff, E. *Fed Proc* **1951**, *10*, 654-659.
 - (14) Franklin, R.; Gosling, R. G. *Nature* **1953**, *171*, 740-741.
 - (15) Franklin, R.; Gosling, R. G. *Nature* **1953**, *172*, 156-157.
 - (16) Furberg, S. *Acta Chem. Scand.* **1952**, *6*, 634-640.
 - (17) Watson, J. D.; Crick, F. H. C. *Nature* **1953**, *171*, 737-738.
 - (18) Pauling, L.; Corey, R. B. *Nature* **1953**, *171*.
 - (19) Pauling, L.; Corey, R. B. *P. Natl. Acad. Sci. USA.* **1953**, *39*, 84.
 - (20) Gulland, J. M.; Jordan, D. O. *Cold Spring Harb. Sym.* **1947**, *12*.
 - (21) Belmont, P.; Constant, J.-F.; Demeunynck, M. *Chem. Soc. Rev.* **2001**, *30*, 70-81.
 - (22) Hunter, C. A. *J. Mol. Biol.* **1993**, *230*, 1025-1054.
 - (23) Kornberg, A.; Baker, T. A. *DNA Replication*; 2nd ed.; W. H. Freeman and Company: New York, **1992**.
 - (24) Waga, S.; Stillman, B. *Annu. Rev. Biochem.* **1998**, *67*, 721-752.
 - (25) Bambara, R. A.; Murante, R. S.; Henricksen, L. A. *J. Biol. Chem.* **1997**, *272*, 4647-4650.
 - (26) Okazaki, R.; Okazaki, T.; Sakabe, K.; Sugimoto, K.; Kainuma, R.; Sugino, A.; Iwatsuki, N. *Cold Spring Harb. Sym.* **1968**, *33*, 129-143.
 - (27) Clancy, S., Online Multimedia, **2008**, DNA Transcription, *Nature Education* 1(1).
 - (28) Clancy, S.; Brown, W., Online Multimedia, **2008**, Translation: DNA to mRNA to protein, *Nature Education* 1(1).

-
- (29) Crick, F. H. C.; S., F. R.; Barnett, L.; Brenner, S.; Watts-Tobin, R. J. *Nature* **1961**, *192*, 1227-1232.
- (30) Crick, F. H. C. *J. Mol. Biol.* **1966**, *19*, 548-555.
- (31) Pray, L. A., Online Multimedia, **2008**, DNA Replication and causes of mutation, *Nature Education* 1(1).
- (32) Aboul-Ela, F.; Koh, D.; Tinoco, I. J.; Martin, F. H. *Nucleic Acids Res.* **1985**, *13*.
- (33) Claverys, J.-P.; Mejean, V.; Gasc, A.-M.; Sicard, A. M. *P. Natl. Acad. Sci. USA.* **1983**, *80*, 5956-5960.
- (34) Dohet, C.; Wagner, R.; Radman, M. *P. Natl. Acad. Sci. USA.* **1985**, *82*, 503-505.
- (35) Baltimore, D. *Nature* **1970**, *226*, 1209-1211.
- (36) Temin, H. M.; Mizutani, S. *Nature* **1970**, *226*, 1211-1213.
- (37) Broder, S. *Antiviral Res.* **2010**, *85*, 1-38.
- (38) Broder, S. *Med. Res. Rev.* **1990**, *10*, 419-439.
- (39) Bennett, C. F.; Swayze, E. E. *Annu. Rev. Pharmacol.* **2010**, *50*, 259-293.
- (40) Mosei, H. E.; Dervan, P. B. *Science* **1987**, *238*, 645-650.
- (41) Le Doan, T.; Perrouault, L.; Praseuth, D.; Habhou, N.; Decout, J. L.; Thuong, N. T.; Lhomme, J.; Helene, C. *Nucleic Acids Res.* **1987**, *15*, 7749-7760.
- (42) Soyfer, V. N.; Potaman, V. N. *Triple Helical Nucleic Acids*; Springer-Verlag New York, Inc, **1996**.
- (43) Praseuth, D.; Guieysse, A. L.; Helene, C. *Biochim. Biophys. Acta* **1999**, *1489*, 181-206.
- (44) Simon, P.; Cannata, F.; Concordet, J.-P.; Giovannangeli, C. *Biochimie* **2008**, *90*, 1109-1116.
- (45) Jain, A.; Wang, G.; Vasquez, K. M. *Biochimie* **2008**, *90*, 1117-1130.
- (46) Frank-Kamenetskii, M. D. *Annu. Rev. Biochem.* **1995**, *64*, 65-95.
- (47) Felsenfeld, G.; Rich, A. *Biochim. Biophys. Acta* **1957**, *26*, 457-468.
- (48) Felsenfeld, G.; Davis, D. R.; Rich, A. *J. Am. Chem. Soc.* **1957**, *79*, 2023-2024.
- (49) Howard, F. B.; Frazier, J.; Lipsett, M. N.; Mills, T. *Biochem. Biophys. Res. Commun.* **1964**, *17*.
- (50) Morgan, A. R.; Wells, R. D. *J. Mol. Biol.* **1968**, *37*, 63-80.
- (51) Larsen, A.; Weintraub, H. *Cell* **1982**, *29*, 609-622.
- (52) Htun, H.; Dahlberg, J. E. *Science* **1989**, *243*, 1571-1576.
- (53) Mirkin, S. M.; Frank-Kamenetskii, M. D. *Annu. Rev. Biophys. Biomol. Struct.* **1994**, *23*, 541-576.

-
- (54) Christophe, D.; Cabrer, B.; Bacolla, A.; Targovnik, H.; Pohl, V.; Vassart, G. *Nucleic Acids Res.* **1985**, *13*, 5127-5144.
- (55) Lee, J., S.; Johnson, D., A.; Morgan, A. R. *Nucleic Acids Res.* **1979**, *6*, 3073-3091.
- (56) Lyamichev, V. I.; Mirkin, S. M.; Frank-Kamenetskii, M. D. *J. Biomol. Struct. Dyn.* **1986**, *3*, 667-669.
- (57) Van Dyke, M. W. *DNA Conformation and Transcription*; Landes Bioscience, **2005**.
- (58) Lyamichev, V. I.; Mirkin, S. M.; Frank-Kamenetskii, M. D. *J. Biomol. Struct. Dyn.* **1985**, *3*, 327-338.
- (59) Hoogsteen, K. *Acta Crystallogr. Sect. A: Found. Crystallogr.* **1959**, *12*, 822-823.
- (60) Beal, P. A.; Dervan, P. B. *Science* **1991**, *251*, 1360-1363.
- (61) Rusling, D. A.; Brown, T.; Fox, K. R. In *Sequence-specific DNA Binding Agents*; Waring, M. J., Ed.; RSC Publishing: Cambridge, UK, 2006, p 1-28.
- (62) Gowers, D. M.; Bijapur, J.; Brown, T.; Fox, K. R. *Biochemistry* **1999**, *38*, 13747-13758.
- (63) Doronina, S. O.; Behr, J.-P. *Chem. Soc. Rev.* **1997**, *26*, 63-71.
- (64) Paes, H. M.; Fox, K. R. *Nucleic Acids Res.* **1997**, *25*, 3269-3274.
- (65) Shimizu, M.; Konishi, A.; Shimada, Y.; Inoue, H.; Ohtsuka, E. *FEBS Lett.* **1992**, *302*, 155-158.
- (66) Sarhan, S.; Seiler, N. *Biol. Chem. Hoppe-Seyler* **1989**, *370*, 1279-1284.
- (67) Pegg, A. E. *Cancer Res.* **1988**, *48*, 759-774.
- (68) Ouameur, A. A.; Tajmir-Riahi, H. *J. Biol. Chem.* **2004**, *279*, 42041-42054.
- (69) Tabor, H. *Biochemistry* **1962**, *1*, 496-501.
- (70) Hampel, K. J.; Crosson, P.; Lee, J., S. *Biochemistry* **1991**, *30*, 4455-4459.
- (71) Thomas, T.; Thomas, T. J. *Biochemistry* **1993**, *32*, 14068-14074.
- (72) Malkov, V. A.; Soyfer, V. N.; Frank-Kamenetskii, M. D. *Nucleic Acids Res.* **1992**, *20*, 4889-4895.
- (73) Cheng, A.-J.; Van Dyke, M. W. *Nucleic Acids Res.* **1993**, *21*, 5630-5635.
- (74) Olivas, W. M.; Maher, L. J. *Nucleic Acids Res.* **1995**, *23*, 1936-1941.
- (75) Chan, P. P.; Glazer, P. M. *J. Mol. Med.* **1997**, *75*, 267-282.
- (76) Olivas, W. M.; Maher, L. J.; III *Biochemistry* **1995**, *34*, 278-284.
- (77) Asensio, J. L.; Lane, A. N.; Dhesi, J.; Bergqvist, S.; Brown, T. *J. Mol. Biol.* **1998**, *275*, 811-822.
- (78) Volker, J.; Klump, H. H. *Biochemistry* **1994**, *33*, 13502-13508.

-
- (79) Roberts, R. W.; Crothers, D. M. *P. Natl. Acad. Sci. USA*. **1996**, *93*, 4320-4325.
- (80) Fox, K. R. *Curr. Med. Chem.* **2000**, *7*, 17-37.
- (81) Leitner, D.; Schroder, W.; Weisz, K. *Biochemistry* **2000**, *39*, 5886-5892.
- (82) Vester, B.; Wengel, J. *Biochemistry* **2004**, *43*, 13233-13241.
- (83) Peterson, M.; Wengel, J. *Trends Biotechnol.* **2003**, *21*, 74-81.
- (84) Asensio, J. L.; Carr, R.; Brown, T.; Lane, A. N. *J. Am. Chem. Soc.* **1999**, *121*, 11063-11070.
- (85) Wolfe, S. *Acc. Chem. Res.* **1972**, *5*, 102-111.
- (86) Plavec, J.; Thibaudeau, C.; Chattopadhyaya, J. *J. Am. Chem. Soc.* **1994**, *116*, 6558-6560.
- (87) Escude, C.; Sun, J. S.; Rougee, M.; Garestier, T.; Helene, C. *C.R. Acad. Sci., III-Vie.* **1992**, *315*, 521-525.
- (88) Wang, S.; Kool, E. T. *Nucleic Acids Res.* **1995**, *23*, 1157-1164.
- (89) Sproat, B. S.; Lamond, A. I.; Beijer, B.; Neuner, P.; Ryder, U. *Nucleic Acids Res.* **1989**, *17*, 3373-3386.
- (90) Bolli, M.; Litten, J. C.; Schutz, R.; Leumann, C. J. *Chem. Biol.* **1996**, *3*, 197-206.
- (91) Nielsen, P.; Pfundheller, H. M.; Wengel, J. *Chem. Commun.* **1997**, 825-826.
- (92) Nielsen, P.; Pfundheller, H. M.; Olsen, C. E.; Wengel, J. *J. Chem. Soc., Perkin Trans. I* **1997**, 3423-3433.
- (93) Singh, S. K.; Nielsen, P.; Koshkin, A. A.; Wengel, J. *Chem. Commun.* **1998**, 455-456.
- (94) Torigoe, H.; Hari, Y.; Sekiguchi, M.; Obika, S.; Imanshi, T. *J. Mol. Biol.* **2001**, *276*, 2354-2360.
- (95) Gotfredsen, C. H.; Schultze, P.; Feigon, J. *J. Am. Chem. Soc.* **1998**, *120*, 4281-4289.
- (96) Dempcy, R. O.; Browne, K. A.; Bruice, T. C. *P. Natl. Acad. Sci. USA*. **1995**, *92*, 6097-6101.
- (97) Blasko, A.; Dempcy, R. O.; Minyat, E. E.; Bruice, T. C. *J. Am. Chem. Soc.* **1996**, *118*, 7892-7899.
- (98) Park, M.; Bruice, T. C. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 3247-3251.
- (99) Park, M.; Toporowski, J. W.; Bruice, T. C. *Biorg. Med. Chem.* **2006**, *14*, 1743-1749.
- (100) Dagle, J. M.; Weeks, D. L. *Nucleic Acids Res.* **1996**, *24*, 2143-2149.

-
- (101) Chaturvedi, S.; Horn, T.; Letsinger, R. L. *Nucleic Acids Res.* **1996**, *24*, 2318-2323.
- (102) Michel, T.; Debart, F.; Heitz, F.; Vasseur, J.-J. *ChemBioChem* **2005**, *6*, 1254-1262.
- (103) Cuenoud, B.; Casset, F.; Husken, D.; Natt, F.; Wolf, R. M.; Altmann, K.-H.; Martin, P.; Moser, H. E. *Angew. Chem. Int. Edit.* **1998**, *37*, 1288-1291.
- (104) Atsumi, N.; Ueno, Y.; Kanazaki, M.; Shuto, S.; Matsuda, C. *Biorg. Med. Chem.* **2002**, *10*, 2933-2939.
- (105) Blommers, M. J. J.; Natt, F.; Jahne, W.; Cuenoud, B. *Biochemistry* **1998**, *37*, 17714-17725.
- (106) Bijapur, J.; Keppler, M. D.; Bergqvist, S.; Brown, T. *Nucleic Acids Res.* **1999**, *27*, 1802-1809.
- (107) Nara, H.; Ono, A.; Matsuda, A. *Bioconjugate Chem.* **1995**, *6*, 54-61.
- (108) Rajeev, K. G.; Jadhav, V. R.; Ganesh, K. N. *Nucleic Acids Res.* **1997**, *25*, 4187-4193.
- (109) Sollogoub, M.; Dominguez, B.; Fox, K. R.; Brown, T. *Chem. Commun.* **2000**, 2315-2316.
- (110) Sollogoub, M.; Darby, R. A. J.; Cuenoud, B.; Brown, T.; Fox, K. R. *Biochemistry* **2002**, *41*, 7224-7231.
- (111) Osborne, S. D.; Powers, V. E. C.; Rusling, D. A.; Lack, O.; Fox, K. R.; Brown, T. *Nucleic Acids Res.* **2004**, *32*, 4439-4447.
- (112) Rusling, D. A.; Powers, V. E. C.; Ranasinghe, R. T.; Wang, Y.; Osborne, S. D.; Brown, T.; Fox, K. R. *Nucleic Acids Res.* **2005**, *33*, 3025-3032.
- (113) Miller, P. S.; Dreon, N.; Pulford, S. M.; McParland, K. B. *J. Biol. Chem.* **1980**, *255*, 9659-9665.
- (114) Kibler-Herzog, L.; Kell, B.; Zon, G.; Shinozuka, K.; Mizan, S.; Wilson, W. D. *Nucleic Acids Res.* **1990**, *18*, 3345-3555.
- (115) Debart, F.; Meyer, A.; Vasseur, J.-J.; Rayner, B. *Nucleic Acids Res.* **1998**, *26*, 4551-4556.
- (116) Gryaznov, S. M.; Lloyd, D. H.; Chen, J.-K.; Shultz, R. G.; DeDionisio, L. A.; Ratmeyer, L.; Wilson, W. D. *P. Natl. Acad. Sci. USA.* **1995**, *92*.
- (117) Torigoe, H. *Biochemistry* **2001**, *40*, 1063-1069.
- (118) Torigoe, H.; Maruyama, A. *J. Am. Chem. Soc.* **2005**, *127*, 1705-1710.

-
- (119) Lacroix, L.; Arimondo, P. B.; Takasugi, M.; Helene, C.; Mergny, J.-L. *Biochem. Biophys. Res. Commun.* **2000**, *270*, 363-369.
- (120) Egholm, M.; Buchardt, O.; Nielsen, P. E.; Berg, R. H. *J. Am. Chem. Soc.* **1992**, *114*, 1895-1897.
- (121) Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science* **1991**, *254*, 1497-1500.
- (122) Egholm, M.; Buchardt, O.; Christensen, L.; Behrens, C.; Freier, S. M.; Driver, D. A.; Berg, R. H.; Kim, S. K.; Norden, B.; Nielsen, P. E. *Nature* **1993**, *365*, 566-568.
- (123) Wittung, P.; Nielsen, P.; Norden, B. *J. Am. Chem. Soc.* **1996**, *118*, 7049-7054.
- (124) Nielsen, P.; Christensen, A. T. *J. Am. Chem. Soc.* **1996**, *118*, 2287-2288.
- (125) Wittung, P.; Nielsen, P.; Norden, B. *Biochemistry* **1997**, *36*, 7973-7979.
- (126) Sponer, J.; Leszczynski, J.; Hobza, P. *J. Phys. Chem.* **1996**, *100*, 5590-5596.
- (127) Petersheim, M.; Turner, D. H. *Biochemistry* **1983**, *22*, 256-263.
- (128) Staubli, A. B.; Dervan, P. B. *Nucleic Acids Res.* **1994**, *22*, 2637-2642.
- (129) Michel, J.; Toulme, J.-J.; Vercauteren, J.; Moreau, S. *Nucleic Acids Res.* **1996**, *24*, 1127-1135.
- (130) Godde, F.; Toulme, J.-J.; Moreau, S. *Biochemistry* **1998**, *37*, 13765-13775.
- (131) Xodo, L. E.; Manzini, G.; Quadrifoglio, F.; van der Marel, G. A.; van Boom, J. H. *Nucleic Acids Res.* **1991**, *19*, 5625-5631.
- (132) Colocci, N.; Dervan, P. B. *J. Am. Chem. Soc.* **1994**, *116*, 785-786.
- (133) Phipps, A. K.; Tarkoy, M.; Schultz, P.; Feigon, J. *Biochemistry* **1998**, *37*, 5820-5830.
- (134) Lacroix, L.; Lacoste, J.; Reddoch, J. F.; Mergny, J.-L.; Levy, D. D.; Seidman, M. M.; Matteucci, M.; Glazer, P. M. *Biochemistry* **1999**, *38*, 1893-1901.
- (135) Froehler, B. C.; Wadwani, S.; Terhorst, T. J.; Gerrard, S., R.. *Tetrahedron Lett.* **1992**, *33*, 5307-5310.
- (136) Mergny, J.-L.; Duval-Valentin, G.; Nguyen, C. H.; Perrouault, L.; Faucon, B.; Rougee, M.; Montenay-Garestier, T.; Bisagni, E.; Helene, C. *Science* **1992**, *256*, 1681-1684.
- (137) Asseline, U.; Hau, J.-F.; Czernecki, S.; Le Diguarher, T.; Perlat, M.-C.; Valery, J.-M.; Thuong, N. T. *Nucleic Acids Res.* **1991**, *19*, 4067-4074.
- (138) Silver, G. C.; Sun, J.-S.; Nguyen, C. H.; Boutorine, A. S.; Bisagni, E.; Helene, C. *J. Am. Chem. Soc.* **1997**, *119*, 263-268.

-
- (139) Robles, J.; McLaughlin, L. W. *J. Am. Chem. Soc.* **1997**, *119*, 6014-6021.
- (140) Capobianco, M. L.; De Champdore, M.; Arcamone, F.; Garbesi, A.; Guianvarc'h, D.; Arimondo, P. B. *Bior. Med. Chem.* **2005**, *13*, 3209-3218.
- (141) Kukreti, S.; Sun, J.-S.; Garestier, T.; Helene, C. *Nucleic Acids Res.* **1997**, *25*, 4264-4270.
- (142) Takasugi, M.; Guendouz, A.; Chassignol, M.; Decout, J. L.; Lhomme, J.; Thuong, N. T.; Helene, C. *P. Natl. Acad. Sci. USA.* **1991**, *88*, 5602-5606.
- (143) Lee, J., S.; Woodsworth, M. L.; Latimer, L. J. P.; Morgan, A. R. *Nucleic Acids Res.* **1984**, *12*, 6603-6614.
- (144) Povsic, T. J.; Dervan, P. B. *J. Am. Chem. Soc.* **1989**, *111*, 3059-3061.
- (145) Koh, J. S.; Dervan, P. B. *J. Am. Chem. Soc.* **1992**, *114*, 1470-1478.
- (146) Singleton, S. F.; Dervan, P. B. *Biochemistry* **1992**, *31*, 10995-11003.
- (147) Seela, F.; Shaikh, K. I. *Org. Biomol. Chem.* **2006**, *4*, 3993-4004.
- (148) Ono, A.; Ts'o, P. O. P.; Kan, L. *J. Am. Chem. Soc.* **1991**, *113*, 4032-4033.
- (149) Ono, A.; Ts'o, P. O. P.; Kan, L. *J. Org. Chem.* **1992**, *57*, 3225-3230.
- (150) Mayer, A.; Leumann, C. J. *Nucleos. Nucleot. Nucl.* **2003**, *22*, 1919-1925.
- (151) Egholm, M.; Christensen, L.; Dueholm, K. L.; Buchardt, O.; Coull, J.; Nielsen, P. *Nucleic Acids Res.* **1995**, *23*, 217-222.
- (152) Berressem, R.; Engels, J. W. *Nucleic Acids Res.* **1995**, *23*, 3465-3472.
- (153) Xiang, G.; Soussou, W.; McLaughlin, L. W. *J. Am. Chem. Soc.* **1994**, *116*, 11155-11156.
- (154) Xiang, G.; Bogacki, R.; McLaughlin, L. W. *Nucleic Acids Res.* **1996**, *24*, 1963-1970.
- (155) Xiang, G.; McLaughlin, L. W. *Tetrahedron* **1998**, *54*, 375-392.
- (156) Bates, P. J.; Laughton, C. A.; Jenkins, T. C.; Capaldi, D. C.; Roselt, P. D.; Reese, C. B.; Neidle, S. *Nucleic Acids Res.* **1996**, *24*, 4176-4184.
- (157) Cassidy, S. A.; Slickers, P.; Trent, J. O.; Capaldi, D. C.; Roselt, P. D.; Reese, C. B.; Neidle, S.; Fox, K. R. *Nucleic Acids Res.* **1997**, *25*, 4891-4898.
- (158) Hildbrand, S.; Blaser, A.; Parel, S. P.; Leumann, C. J. *J. Am. Chem. Soc.* **1997**, *119*, 5499-5511.
- (159) Miller, P. S.; Bhan, P.; Cushman, C. D.; Trapane, T. L. *Biochemistry* **1992**, *31*, 6788-6793.
- (160) Krawczyk, S. H.; Milligan, J. F.; Wadwani, S.; Moulds, C.; Froehler, B. C.; Matteucci, M. D. *P. Natl. Acad. Sci. USA.* **1992**, *89*, 3761-3764.

-
- (161) Hunziker, J.; Priestly, E. S.; Brunar, H.; Dervan, P. B. *J. Am. Chem. Soc.* **1995**, *117*, 2661-2662.
- (162) Koshlap, K. M.; Schultze, P.; Brunar, H.; Dervan, P. B.; Feigon, J. *Biochemistry* **1997**, *36*, 2659-2668.
- (163) Marfurt, J.; Leumann, C. *Angew. Chem. Int. Ed.* **1998**, *37*, 175-177.
- (164) Marfurt, J.; Parel, S. P.; Leumann, C. *J. Nucleic Acids Res.* **1997**, *25*, 1875-1882.
- (165) Gowers, D. M.; Fox, K. R. *Nucleic Acids Res.* **1999**, *27*, 1569-1577.
- (166) Griffin, L. C.; Dervan, P. B. *Science* **1989**, *245*, 967-971.
- (167) Yoon, K.; Hobbs, C. A.; Koch, J.; Sardaro, M.; Kutny, R.; Weis, A. L. *P. Natl. Acad. Sci. USA.* **1992**, *89*, 3840-3844.
- (168) Gowers, D. M.; Fox, K. R. *Nucleic Acids Res.* **1997**, *25*, 3787-3794.
- (169) Radhakrishnan, I.; Gao, X. L.; de los Santos, C.; Live, D.; Patel, D. J. *Biochemistry* **1991**, *30*, 9022-9030.
- (170) Radhakrishnan, I.; Patel, D. J. *J. Mol. Biol.* **1994**, *241*, 600-619.
- (171) Horne, D. A.; Dervan, P. B. *Nucleic Acids Res.* **1991**, *19*, 4963-4965.
- (172) Loakes, D. *Nucleic Acids Res.* **2001**, *29*, 2437-2447.
- (173) Durland, R. H.; Rao, T. S.; Revankar, G. R.; Tinsley, J. H.; Myrick, M. A.; Seth, D. M.; Rayford, J.; Singh, P.; Jayaraman, K. *Nucleic Acids Res.* **1994**, *22*, 3233-3240.
- (174) Prevot-Halter, I.; Leumann, C. J. *Bioorg. Med. Chem. Lett.* **1999**, *9*, 2657-2660.
- (175) Buchini, S.; Leumann, C. J. *Tetrahedron Lett.* **2003**, *44*, 5065-5068.
- (176) Buchini, S.; Leumann, C. J. *Angew. Chem.* **2004**, *116*, 4015-4018.
- (177) Ranasinghe, R. T.; Rusling, D. A.; Powers, V. E. C.; Fox, K. R.; Brown, T. *Chem. Commun.* **2005**, 2555-2557.
- (178) Chen, D. L.; McLaughlin, L. W. *J. Org. Chem.* **2000**, *65*, 7468-7474.
- (179) Obika, S.; Hari, Y.; Sekiguchi, M.; Imanshi, T. *Angew. Chem. Int. Ed.* **2001**, *40*, 2079-2081.
- (180) Eldrup, A. B.; Dahl, O.; Nielsen, P. *J. Am. Chem. Soc.* **1997**, *119*, 11116-11117.
- (181) Olsen, A. G.; Dahl, O.; Nielsen, P. *Nucleos. Nucleot. Nucl.* **2003**, *22*, 1331-1333.
- (182) Olsen, A. G.; Dahl, O.; Nielsen, P. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 1551-1554.
- (183) Griffin, L. C.; Kiessling, L. L.; Beal, P. A.; Gillespie, P.; Dervan, P. B. *J. Am. Chem. Soc.* **1992**, *114*, 7976-7982.

-
- (184) Koshlap, K. M.; Gillespie, P.; Dervan, P. B.; Feigon, J. *J. Am. Chem. Soc.* **1993**, *115*, 7908-7909.
- (185) Lehmann, T. E.; Greenberg, W. A.; Liberles, D. A.; Wada, C. K.; Dervan, P. B. *Helv. Chim. Acta* **1997**, *80*, 2002-2022.
- (186) Huang, C.-Y.; Cushman, C. D.; Miller, P. S. *J. Org. Chem.* **1993**, *58*, 5048-5049.
- (187) Huang, C.-Y.; Miller, P. S. *J. Am. Chem. Soc.* **1993**, *115*, 10456-10457.
- (188) Huang, C.-Y.; G., B.; Miller, P. S. *Nucleic Acids Res.* **1996**, *24*, 2606-2613.
- (189) Guianvarc'h, D.; Benhida, R.; Fourrey, J.-L.; Maurisse, R.; Sun, J.-S. *Chem. Commun.* **2001**, 1814-1815.
- (190) Guianvarc'h, D.; Fourrey, J.-L.; Maurisse, R.; Sun, J.-S.; Benhida, R. *Bioorg. Med. Chem. Lett.* **2003**, *11*, 2751-2759.
- (191) Guianvarc'h, D.; Fourrey, J.-L.; Maurisse, R.; Sun, J.-S.; Benhida, R. *Org. Lett.* **2002**, *4*, 4209-4212.
- (192) Wang, Y.; Rusling, D. A.; Powers, V. E. C.; Lack, O.; Osborne, S. D.; Fox, K. R.; Brown, T. *Biochemistry* **2005**, *44*, 5884-5892.
- (193) Craynest, N. V.; Guianvarc'h, D.; Peyron, C.; Benhida, R. *Tetrahedron Lett.* **2004**, *45*, 6243-6247.
- (194) Besch, R.; Giovannangeli, C.; Schuh, T.; Kammerbauer, C.; Degitz, K. *J. Mol. Biol.* **2004**, *341*, 979-989.
- (195) Gore, M. G. *Spectrophotometry and spectrofluorimetry*; Oxford University Press: Oxford, **2000**.
- (196) Boutorine, A. S.; Escude, C. In *Current Protocols in Nucleic Acid Chemistry*; John Wiley and Sons, Inc.: Paris, 2007.
- (197) Tinoco, I. J. *J. Am. Chem. Soc.* **1959**, *82*, 4785-4790.
- (198) Rhodes, W. *J. Am. Chem. Soc.* **1961**, *83*, 3609-3617.
- (199) Hampshire, A. J.; Rusling, D. A.; Broughton-Head, V. J.; Fox, K. R. *Methods* **2007**, *42*, 128-140.
- (200) Revzin, A. *Footprinting of Nucleic Acid-Protein Complexes*; Academic Press Inc, **1993**.
- (201) Powers, V. E. C., Thesis, *Synthesis and studies of nucleoside analogues in triplex forming oligonucleotides*, University of Southampton, **2004**.
- (202) Reese, C. B.; Wu, Q. *Org. Biomol. Chem.* **2003**, *1*, 3160-3172.
- (203) Cimino, G. D.; Gamper, H. B.; Isaacs, S. T.; Hearst, J. E. *Annu. Rev. Biochem.* **1985**, *54*, 1151-1193.

-
- (204) Watts, J. *Professional Nurse* **1999**, *14*, 623-626.
- (205) Davison, S.; Poyner, T.; Barker, J. *Psoriasis* **2000**, *Blackwell Science*.
- (206) Donaldson, L.; Stewart Douglas, W. *The Practitioner* **1997**, *241*, 66-72.
- (207) Galloway, G. A.; Lawson, G. B. *Dermatology Nursing* **1995**, *7*, 348-351.
- (208) Francis, M. M.; Mellem, J. E.; Maricq, A. V. *Trends. Neurosci.* **2003**, *26*, 90-99.
- (209) Brenner, S. *Genetics* **1974**, *77*, 71-94.
- (210) Maeda, I.; Kohara, Y.; Yamamoto, M.; Sugimoto, A. *Curr. Biol.* **2001**, *11*, 171-176.
- (211) Hildbrand, S.; Leumann, C. *Angew. Chem. Int. Edit.* **1996**, *35*, 1968-1970.
- (212) Bates, P. J.; Laughton, C. A.; Jenkins, T. C.; Capaldi, D. C.; Roselt, P. D.; Reese, C. B.; Neidle, S. *Nucleic Acids Res.* **1996**, *24*, 4176-4184.
- (213) Shapiro, R.; Chambers, R. W. *J. Am. Chem. Soc.* **1961**, *83*, 3920-3921.
- (214) Adamo, M. F. A.; Adlington, R. M.; Baldwin, J. E.; Day, A. L. *Tetrahedron* **2004**, *60*, 841-849.
- (215) Ohrui, H.; Jones, G. H.; Moffatt, J. G.; Maddox, M. L.; Christensen, A. T.; Byram, S. K. *J. Am. Chem. Soc.* **1975**, *97*.
- (216) Ohno, M.; Ito, Y.; Arita, M.; Shibata, T.; Adachi, K.; Sawai, H. *Tetrahedron* **1984**, *40*, 145-152.
- (217) Kozikowski, A. P.; Cheng, X.-M. *J. Chem. Soc., Chem. Commun.* **1987**, 680-683.
- (218) Singh, I.; Seitz, O. *Org. Lett.* **2006**, *8*, 4319-4322.
- (219) Brown, D. M.; Ogden, R. C. *J. Chem. Soc. Perk. T. 1* **1981**, 723-725.
- (220) Harusawa, S.; Murai, Y.; Moriyama, H.; Imazu, T.; Ohishi, H.; Yoneda, R.; Kurihara, T. *J. Org. Chem.* **1996**, *61*, 4405-4411.
- (221) Harusawa, S.; Matsuda, C.; Araki, L.; Kurihara, T. *Synthesis* **2006**, 793-798.
- (222) Kalvoda, L.; Farkas, J.; Sorm, F. *Tetrahedron Lett.* **1970**, *26*, 2297-2300.
- (223) Krohn, K.; Heins, H.; Wielcken, K. *J. Med. Chem.* **1992**, *35*, 511-517.
- (224) Raboisson, P.; Baurand, A.; Cazenave, J. P.; Gachet, C.; Schultz, D.; Spiess, B.; Bourguignon, J. J. *J. Org. Chem.* **2002**, *67*, 8063-8071.
- (225) Watanabe, K. A. In *Chemistry of Nucleosides and Nucleotides*; Townsend, L. B., Ed.; Plenum Press, New York: 1994; Vol. 3, p 421-435.
- (226) Postema, M. H. D. In *C-Glycoside Synthesis - New Directions in Organic and Biological Chemistry*; CRC Press: Boca Raton, FL USA, 1995, p 303.
- (227) Stambasky, J.; Hocek, M.; Kocovsky, P. *Chem. Rev* **2009**, *109*, 6729-6764.
- (228) Lerch, U.; Burdon, M. G.; Moffatt, J. G. *J. Org. Chem.* **1971**, *36*, 1507-1513.

-
- (229) Brown, D. M.; Burdon, M. G.; Slatcher, R. P. *Chem. Commun.* **1965**, 5, 77-78.
- (230) Cameron, M. A.; Cush, S. B.; Hammer, R. P. *J. Org. Chem.* **1997**, 62, 9065-9069.
- (231) Barton, D. H. R.; McCombie, S. W. *J. Chem. Soc., Perkin Trans. 1* **1975**, 1574-1585.
- (232) Hurd, C. D.; Bonner, W. A. *J. Am. Chem. Soc.* **1945**, 67, 1972-1977.
- (233) Hoffer, M. *Chem. Ber.* **1960**, 93, 2777-2781.
- (234) Chu, C. K.; Wempen, I.; Watanabe, K. A.; Fox, J. J. *J. Org. Chem.* **1976**, 41, 2793-2797.
- (235) Maeba, I.; Iwata, K.; Usami, F.; Furukawa, H. *J. Org. Chem.* **1983**, 48, 2998-3002.
- (236) Kraus, G. A.; Molina, M. T. *J. Org. Chem.* **1988**, 53, 752-753.
- (237) Hanessian, S.; Machaalani, R. *Tetrahedron Lett.* **2003**, 44, 8321-8323.
- (238) Arai, I.; Daves, G. D. *J. Am. Chem. Soc.* **1978**, 100, 287-288.
- (239) Walker, J. A.; Liu, W.; Wise, D. S.; Drach, J. C.; Townsend, L. B. *Tetrahedron Lett.* **1998**, 41, 1236-1241.
- (240) Jiang, Y. L.; Stivers, J. T. *Tetrahedron Lett.* **2003**, 44, 85-88.
- (241) Jiang, Y. L.; Stivers, J. T. *Tetrahedron Lett.* **2003**, 44, 4051-4055.
- (242) Hocek, M.; Pohl, R.; Klepetarova, B. *Eur. J. Org. Chem.* **2005**, 4525-4528.
- (243) Urban, M.; Pohl, R.; Klepetarova, B.; Hocek, M. *J. Org. Chem.* **2005**, 71, 7322-7328.
- (244) Joubert, N.; Pohl, R.; Klepetarova, B.; Hocek, M. *J. Org. Chem.* **2007**, 72, 6797-6805.
- (245) Barta, J.; Pohl, R.; Klepetarova, B.; Ernsting, N. P.; Hocek, M. *J. Org. Chem.* **2008**, 73, 3798-3806.
- (246) Joubert, N.; Urban, M.; Pohl, R.; Hocek, M. *Synthesis* **2008**, 12, 1918-1932.
- (247) Barker, R.; Fletcher, H. G. *J. Org. Chem.* **1961**, 26, 4605-4609.
- (248) Ley, S. V.; Norman, J.; Griffith, W. P.; Marsden, S. P. *Synthesis* **1994**, 639-666.
- (249) Timpe, W.; Dax, K.; Wolf, N.; Weidmann, H. *Carbohydr. Res.* **1975**, 39, 53-60.
- (250) Masaoki, I.; Eckstein, F. *J. Org. Chem.* **1979**, 44, 2039-2041.
- (251) Pyne, S. G. *Tetrahedron Lett.* **1987**, 28, 4737-4740.
- (252) Goody, R. S.; Walker, R. T. *Tetrahedron Lett.* **1967**, 8, 289-291.
- (253) Takahashi, T.; Hirose, Y.; Iwamoto, H.; Doi, T. *J. Org. Chem.* **1998**, 63, 5742-5743.

-
- (254) Greene, T. W.; Wuts, P. G. M. *Protective groups in organic synthesis*; 3rd ed.; John Wiley & Sons, Inc., **1999**.
- (255) Seela, F.; Peng, X. *Synthesis* **2004**, 8, 1203-1210.
- (256) Wuts, P. G. M.; Greene, T. W. *Protective Groups in Organic Synthesis*; Fourth ed.; John Wiley & Sons Inc, **2007**.
- (257) Heathcock, C. H.; Ratcliffe, R. *J. Am. Chem. Soc.* **1971**, 93, 1746-1757.
- (258) Witte, J. F.; McClard, R. W. *Bioorg. Chem.* **1996**, 24, 29-38.
- (259) Ito, K.; Takasawa, T.; Ohba, Y. *Synth. Commun.* **2002**, 32 3839-3849.
- (260) Inaba, T.; Umezawa, I.; Yuasa, M.; Inoue, T.; Mihashi, S.; Itokawa, H.; Ogura, K. *J. Org. Chem.* **1987**, 52.
- (261) Yokoyama, M.; Ikuma, T.; Obars, N.; Togo, H. *J. Chem. Soc. Perk. T. I* **1990**, 12, 3243-3247.
- (262) Wierenge, W.; Skulnick, H. I. *Carbohydr. Res.* **1981**, 90, 41-52.
- (263) Tsunoda, T.; Otsuka, J.; Yamamiya, Y.; Ito, S. *Chem. Lett.* **1994**, 539-542.
- (264) Tsunoda, T.; Nagaku, M.; Nagino, C.; Kawamura, Y.; Ozaki, F.; Hioki, H.; Ito, S. *Tetrahedron Lett.* **1995**, 36, 2531-2534.
- (265) Dodge, J., A.; Nissen, J., S.; Presnell, M. *Org. Synth.* **1996**, 9, 607.
- (266) Lu, W.; Sengupta, S.; Peterson, J., L.; Akhmedov, N., G.; Shi, X. *J. Org. Chem.* **2007**, 72, 5012-5015.
- (267) Nitthyanandhan, J.; Jayaraman, N. *Tetrahedron* **2005**, 61, 11184-11191.
- (268) Srivastava, P. C.; Robbins, R. K.; Takusagawa, F.; Berman, H. M. *J. Heterocycl. Chem.* **1981**, 18, 1659-1662.
- (269) Yokoyama, M.; Toyoshima, A.; Akiba, T.; Togo, H. *Chem. Lett.* **1994**, 23, 265.
- (270) Bellamy, A. J.; MacCuish, A. *Propell. Explos. Pyrot.* **2007**, 32, 20-31.
- (271) ElAmin., B.; Anantharamaiah., G. M.; Royer., G. P.; Means., G. E. *J. Org. Chem.* **1979**, 44, 3442.
- (272) Liu, H.; Yip, J. *Tetrahedron Lett.* **1997**, 38, 2253-2256.
- (273) Alonso, E.; Ramon, D. J.; Yus, M. *Tetrahedron Lett.* **1997**, 53, 14355-14368.
- (274) Grobelny, Z. *Eur. J. Org. Chem.* **2004**, 2973-2982.
- (275) El Sagheer, A. H.; Brown, T. In *Current Protocols in Nucleic Acid Chemistry*; John Wiley & Sons, Inc.: 2008.
- (276) Tietze, L. F. e. a. *Chem. Eur. J.* **1996**, 2, 139-148.
- (277) Takase, M.; Morikawa, T.; Abe, H.; Inouye, T. *Org. Lett.* **2003**, 5, 625-628.
- (278) Majumdar, K. C.; Mondal, S. *Tetrahedron Lett.* **2007**, 48, 6951-6953.

-
- (279) Myers, A., G.; Gleason, J., L.; Yoon, T.; Kung, D., W. *J. Am. Chem. Soc.* **1997**, *119*, 656-673.
- (280) Debaene, F.; Da Silva, J. A.; Pianowski, Z.; Duran, F. J.; Winssinger, N. *Tetrahedron* **2007**, *63*, 6577-6586.
- (281) Li, H.; Qiu, Y.; Moyroud, E.; Kishi, Y. *Angew. Chem. Int. Ed.* **2001**, *40*, 1471-1475.
- (282) Lou, C.; Xiao, Q.; Tailor, R. R.; Ben Gaied, N.; Gale, N.; Light, M. E.; Fox, K. R.; Brown, T. *Med. Chem. Commun.* **2011**, DOI 10.1039/c1md00068c.
- (283) Vincent, S.; Mioskowski, C.; Lebeau, L. *J. Org. Chem.* **1999**, *64*, 991-997.
- (284) Puri, N.; Majumdar, K. C.; Cuenoud, B.; Natt, F.; Martin, P.; Boyd, A.; Miller, P. S.; Seidman, M. M. *J. Biol. Chem.* **2001**, *276*, 28991-28998.
- (285) Roberts, R. W.; Crothers, D. M. *Science* **1992**, *258*, 1463-1466.
- (286) Puri, N.; Majumdar, A.; Cuenoud, B.; Natt, F.; Martin, P.; Boyd, A.; Miller, P. S.; Seidman, M. M. *Biochemistry* **2002**, *41*, 7716-7724.
- (287) Rajagopal, P.; Feigon, J. *Biochemistry* **1989**, *28*, 7859-7870.
- (288) Umemoto, K.; Sarma, M. H.; Gupta, G.; Luo, J.; Sarma, R. H. *J. Am. Chem. Soc.* **1990**, *112*, 4539-4545.
- (289) Arnott, A.; Chandrasekaran, R.; Hukins, D. W. L.; Smith, P. J. C.; Watts, L. *J. Mol. Biol.* **1974**, *88*, 523-524.
- (290) Buchini, S.; Leumann, C. J. *Angew. Chem. Int. Edit.* **2004**, *43*, 3925-3928.
- (291) Cuenoud, B.; Casset, F.; Husken, D.; Natt, F.; Wolf, R. M.; Altmann, K.; Martin, P.; Moser, H. E. *Angew. Chem. Int. Ed.* **1998**, *37*.
- (292) Jin, S.; Miduturu, C. V.; McKinney, D. C.; Silverman, S. K. *J. Org. Chem.* **2005**, *70*, 4284-4299.
- (293) Juliano, R.; Alam, M. R.; Dixit, V.; Kang, H. *Nucleic Acids Res.* **2008**, *36*, 4158-4171.
- (294) Allerson, C. R.; Sioufi, N.; Jarres, R.; Prakash, T. P.; Naik, N.; Berdeja, A.; Wanders, L.; Griffey, R. H.; Swayze, E. E.; Bhat, B. *J. Med. Chem.* **2005**, *48*, 901-904.
- (295) Buhr, C. A.; Matteucci, M.; *U.S. Patent Number 5466786*, Gilead Sciences (Foster City, CA): United States, **1995**.
- (296) Li, H.; Broughton-Head, V. J.; Peng, G.; Powers, V. E. C.; Ovens, M. J.; Fox, K. R.; Brown, T. *Bioconjugate Chem.* **2006**, *17*.

-
- (297) Takasugi, M.; Guendouz, A.; Chassignol, M.; Decout, J. L.; Lhomme, J.; Thuong, N. T.; Helene, C. *P. Natl. Acad. Sci. USA*. **1991**, *88*, 5602-5606.
- (298) Caplen, N. J.; Parrish, S.; Imani, F.; Fire, A.; Morgan, R. A. *P. Natl. Acad. Sci. USA*. **2001**, *98*, 9742-9747.
- (299) Yang, N.-S.; Burkholder, J.; Roberts, B.; Martinell, B.; McCabe, D. *P. Natl. Acad. Sci. USA*. **1990**, *87*, 9568-9572.
- (300) Rols, M.-P.; Deltell, C.; Golzio, M.; Dumond, P.; Cros, S.; Teissie, J. *Nat. Biotechnol.* **1998**, *16*, 168-171.
- (301) Uhlmann, E.; Peyman, A. *Chem. Rev.* **1990**, *90*, 543-584.
- (302) Li, S.; Huang, L. *Gene Ther.* **2000**, *7*, 31-34.
- (303) Gottlieb, H. E.; Kotlyar, V.; Nudelman, A. *J. Org. Chem.* **1997**, *62*, 7512-7515.
- (304) Boyd, E.; Chavda, S.; Eames, J.; Yohannes, Y. *Tet: Asym.* **2007**, *18*, 476-482.
- (305) Schwarz, F. P.; Robinson, S.; Butler, J. M. *Nucl. Acids Res.* **1999**, *27*, 4792-4800.
- (306) Langley, G. J.; Herniman, J. M.; Davies, N. L.; Brown, T. *Rapid Commun. Mass Sp.* **1999**, *13*, 1717-1723.
- (307) Beynon, R. J.; Easterby, J. S. *Buffer Solutions: The Basics*; Taylor & Francis Ltd: Oxford, **2003**.

CHAPTER 7

Appendix

7.1 PUBLICATIONS

1. Lou, C.; Xiao, Q.; Tailor, R. R.; Ben Gaied, N.; Gale, N.; Light, M. E.; Fox, K. R.; Brown, T. *Med. Chem. Commun.* **2011**, DOI 10.1039/c1md00068c.